Psychological Bulletin

CONTENTS

Electrical Stimulation of the Brain and the Psychophysiology of Learning and Motivation ..... H. PHILIP ZEIGLER 363

Laboratory Studies of Behavior Without Awareness ... JOE K. ADAMS 383

On the Application of Genetic Expectancies as Age-Specific Base Rates in the Study of Human Behavior Disorders ..... JOHN S. PEARSON AND IRENE B. KLEV 406

The -ilies That Plague Elementary Statistics ..... HORACE B. ENGLISH 421

A Note Concerning Kendall's Tau ..... DESMOND S. CARTWRIGHT 423

Comment on a Distribution-Free Factorial-Design Analysis ..... FRED D. SHEFFIELD 426

The Statistical Concepts of Confidence and Significance ..... ROBERT E. CHANDLER 429

Published Bimonthly by the American Psychological Association

VOL. 54, NO. 5 SEPTEMBER, 1957

WAYNE DENNIS, Editor
Brooklyn College

Consulting Editors

LAUNOR F. CARTER
RAND Corporation
Santa Monica, California

EDWARD GIRDEN
Brooklyn College

VICTOR C. RANEY
University of Colorado

ROBERT [illegible]
Teachers College, Columbia University

BENTON J. [illegible]
Northwestern University

S. RAINS [illegible]
Life Insurance Agency Management Association

ARTHUR C. HOFFMAN, Managing Editor

HELEN ORR, Assistant Managing Editor

Editorial Staff: FRANCES H. CLARK, BARBARA CUMMINGS, SARAH J. DOYLE, SARAH WOMACK

The Psychological Bulletin contains evaluative reviews of research literature and articles on research methodology in psychology. This journal does not publish *reports of original research or original theoretical articles.*

Manuscripts should be sent to Wayne Dennis, Department of Psychology, Brooklyn College, Brooklyn 10, New York.

Preparation of articles for publication. Authors are strongly advised to follow the general directions given in the "Publication Manual of the American Psychological Association" (*Psychological Bulletin*, 1952, 49 [No. 4, Part 2], 389–449). Special attention should be given to the section on the preparation of the references (pp. 432–449), since this is a particular source of difficulty in long reviews of research literature. *All copy must be double spaced, including the references.* All manuscripts should be submitted *in duplicate.* Original figures are prepared for publication; duplicate figures may be photographs or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail.

Reprints. Fifty free offprints are given to contributors of articles and notes. Authors of early publication articles receive no gratis offprints.

Communications—including subscriptions, orders of back issues, and changes of address—should be addressed to the American Psychological Association, 1333 Sixteenth Street N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

Annual subscription: $8.00 (Foreign $8.50). Single copies, $1.50.

PUBLISHED BIMONTHLY BY:

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

Menasha, Wisconsin

and 1333 Sixteenth Street N.W., Washington 6, D.C.

Entered as second class mail matter at the post office at Washington, D.C., under the act of March 3, 1879. Additional entry at the post office at Menasha, Wisconsin. Acceptance for mailing at special rate of postage under the provisions of Sec. 34.40 Par. (D) provided for in Section 3.24, act of February 28, 1925, authorized August 6, 1947. Printed in U.S.A.

Copyright, 1957, by The American Psychological Association, Inc.

VOL. 54, NO. 5 SEPTEMBER, 1957

Psychological Bulletin

ELECTRICAL STIMULATION OF THE BRAIN AND THE PSYCHOPHYSIOLOGY OF LEARNING AND MOTIVATION[1]

H. PHILIP ZEIGLER[2]

Departments of Psychology and Neurophysiology,[3] University of Wisconsin

As a physiological technique, electrical stimulation of the brain dates from the classic studies of Fritsch and Hitzig (**24**) on the motor cortex of the dog. The use of this technique in behavioral research, however, is a more recent development. Since the pioneer studies of Loucks (**62**, **63**, **64**), psychologists have evinced sporadic interest in electrical stimulation of the brain as an important research tool in elucidating the neural mechanisms underlying behavior. Recent findings in neurophysiology, and advances in electrophysiological methodology, have led to a resurgence of this interest. Hess's studies of the diencephalon (**41**, **42**, **43**, **44**, **45**, **46**), Magoun's work on reticular activating systems (**70**), and studies of sensory and motor representation by Woolsey (**101**) have been of particular interest to psychologists. Moreover, some of the techniques developed in the course of this research have been adapted for use with human subjects. The studies of Penfield and his associates, on patients undergoing neurosurgery, are classic examples of the application of these techniques to the investigation of human brain function (**88**). The results so obtained have confirmed and extended the findings on cortical localization obtained from lower animals. Studies employing stimulation of the human brain through electrodes implanted in various subcortical structures are currently in progress (**37**, **38**).

The investigations cited above have provided a wealth of factual material regarding neurophysiological processes, but the relation of these processes to behavior remains obscure. As Teuber has pointed out, "The absence of any convincing physiological correlate of learning is the greatest gap in physiological psychology" (**99**). Electrical stimulation of the brain is among the most promising of the available techniques for establishing such correlates. Its greatest advantage over cerebral ablation is that it makes possible the

[1] The following abbreviations have been used in this review: CER (Conditioned Emotional Response), CFR (Conditioned Fear Response), CM (Centrum Medianum nucleus of the thalamus), CNS (Central Nervous System), CR (Conditioned Response), CS (Conditioned Stimulus), EEG (Electroencephalograph), GSR (Galvanic Skin Response), S-S (Sensory-Sensory), S-R (Stimulus Response), UCS (Unconditioned Stimulus), VM (Ventro Medialis nucleus of the thalamus), VPL (Ventro-Postero-Lateralis nucleus of the thalamus).

[2] The author wishes to express his gratitude to Prof. W. J. Brogden, Department of Psychology, and to Drs. K. Akert, R. Benjamin, and W. Welker, Department of Neurophysiology, University of Wisconsin, for their encouragement and helpful criticisms.

[3] This review was completed during the author's tenure as Research Fellow, National Institute of Mental Health.

study of the intact nervous system in the unanesthetized animal under conditions closely approximating those in which the learning process occurs.

This paper is primarily a review of research in which electrical stimulation of the central nervous system is employed in the investigation of neural mechanisms involved in the learning process. Studies of electroconvulsive shock (ECS) and learning have been omitted, because they employ electrical stimulation in a markedly different manner. Although problems of electrophysiological methodology will be briefly discussed wherever relevant, this paper is not primarily concerned with such problems. Readers interested in these problems are referred to publications of Hess (**41**, **46**), Delgado (**17**, **19**), Lilly (**54**, **55**), Heath (**37**), Monnier and Laue (**75**), etc., and to a recent volume on brain stimulation edited by Sheer (**94**).

Intraneural Conditioning

In reviewing a series of studies designed to determine the neural correlates of the conditioned response, Gantt (**27**) suggested the terms intraneural and extraneural conditioning. As opposed to ordinary (extraneural) conditioning, in which the UCS and CS arise outside the central nervous system, intraneural conditioning requires that the UCS and/or CS be produced by direct stimulation of the CNS, and that a CR be established under conditions precluding the possibility of artifactual stimulation. The pioneer studies in this area, those of Loucks, Gantt, and their collaborators, were made feasible by the development of Loucks' technique for brain stimulation of chronic preparations (**61**). The technique utilizes a collodion covered coil, buried just beneath the skin, from which insulated wires lead to an implanted electrode. When the primary of an inductorium is laid over the skin, adjacent to the coil, the induced current is conducted to the electrodes. It was thus possible to substitute direct CNS stimulation for the CS or UCS in a conditioning experiment. (Because of the possibilities of artifactual stimulation (**13**) and certain problems of parametric control, this technique is no longer in general use.)

The results of these early studies of intraneural conditioning in dogs may be summarized as follows:

1. Stimulation of the motor cortex (UCS) paired with an externally presented CS did not produce conditioning. Presentation of food following each trial resulted in effective conditioning (**62**).

2. Stimulation of the visual cortex (CS) paired with an externally presented UCS resulted in the formation of conditioned responses (**63**).

3. Stimulation of the dorsal columns of the spinal cord (UCS) paired with an externally presented CS was ineffective in establishing conditioning. However, stimulation of the dorsal roots (UCS), or an increase in the intensity of dorsal column stimulation resulted in conditioning (**64**).

4. Responses produced by stimulation of the vestibular ganglion (UCS) could be conditioned to a variety of external CS's (**66**).

5. Stimulation of the cerebellum (UCS) produced a variety of responses, some of which could be conditioned to externally presented CS's (**11**).

The studies themselves have been reviewed in detail elsewhere (**34**, **76**). Our concern in the present discussion, therefore, will be with the use which has been made of these findings as experimental support for certain theories of learning.

The results of two of the experiments (**62**, **64**) have been cited by certain theorists as being relevant to one or both of two major theoretical issues. The first of these concerns the role of drive reduction in learning; the second raises the general problem of whether the S-S or S-R paradigm more adequately describes the learning process. Spence (**97**) has cited the study pairing stimulation of the motor cortex with external CS's as support for the drive-reduction viewpoint, since no conditioned responses were established until food was given following each trial (**62**). Supporters of an S-S paradigm have attributed the failure of conditioning in this study to an absence of "contiguity of afferent processes" (**4**, **71**). That is, the assumption is made that the UCR was produced directly through motor stimulation, omitting the afferent input normally provided by an externally presented UCS. The experiment on stimulation of spinal columns vs. roots has been similarly discussed. Spence (**97**) cites this study as evidence for a drive-reduction viewpoint, on the assumption that there was no conditioning until stimulation was intense enough to spread to adjacent pain fibers. Proponents of an S-S paradigm, however, have attributed this failure to "lack of effective sensory relationships of the UCS" (**4**), in a manner similar to their interpretation of the motor stimulation experiment. None of these theorists has discussed the results of intraneural conditioning studies involving cerebellar stimulation (**11**), although in this case conditioning was obtained in some animals, despite by-passing of an external UCS, and absence of external reinforcement.

The relevance to learning theory of interpretations such as these depends, in part, upon the validity of two implicit assumptions. It must first be demonstrated that we are dealing with intraneural and not extraneural conditioning; that is, the CS or UCS must be produced by electrical stimulation of the central nervous system. It is probable that this assumption is not tenable for either experiment. In the study which employed motor stimulation as the UCS and found no conditioning until food was presented following every trial (**62**), there is evidence that the animals were selectively learning through instrumental reward procedures. Loucks attempted to test this hypothesis by presenting the food randomly, rather than following every trial. The results of this control procedure supported the hypothesis of instrumental reward learning, since with this modified procedure there was complete failure to establish conditioned responses (**62**). The second experiment, involving stimulation of the spinal cord, is also not a clear-cut case of intraneural conditioning. No conditioned responses were obtained until the increase in stimulus intensity resulted in a "rough caudal jerk of the leg" (**64**). Loucks points out that it was this rough caudal leg jerk, with its consequent excitation of nociceptive receptors, which was the actual UCS. That is, the UCS was peripheral, not central stimulation.

The relation of these experiments to the S-S, S-R controversy rests on the assumption that the experiments have effectively separated the sensory and motor components of the conditioning process. In the light of current knowledge of the cortical representation of sensory and motor function, this assumption is untenable. There is considerable evidence which indicates that the clear-cut distinc-

tion between cortical sensory and motor areas can no longer be maintained. Sensory representation in the motor cortex (**72**, **100**) and the cerebellum (**96**) has been demonstrated. Moreover, the results of recent investigations "suggest that motor functions are distributed throughout the unanesthetized cerebral cortex in the normal state, and [these results] imply, when correlated with sensory maps, that each small area of the cortex is truly 'sensori-motor', with a preponderance of one or the other function" (**53**).

One final point is worthy of note. None of the hypotheses (S-S vs. S-R, drive reduction vs. sensory integration) have been formulated in a manner that would make them accessible to neurophysiological verification. For this reason, as well as those discussed above, the use of these experiments as support for any given learning theory is probably unjustified. However, these early studies have played an important role in stimulating further research in this area.

In the period following the experiments of Loucks, Gantt, and their associates, work on intraneural conditioning has continued. Reports by Culler (**15**) and Giurgea (**35**) exist only in abstract form, and have never been satisfactorily confirmed. Loucks has recently begun a program of intraneural conditioning studies on the domestic hen (**65**). Conditioned responses are established to electrical stimulation of the brain (CS) and the effects of implanting metal barriers in the vicinity of the electrodes are then investigated. The results indicate that barriers perpendicular to the surface of the brain appear to have negligible effects upon the establishment or retention of conditioned responses.

In a paper concerned primarily with methodological problems in intraneural conditioning, Doty et al. report such conditioning in the cat (**23**). These investigators found that conditioned responses may be established to stimulation of the dura alone and have developed procedures for the control or elimination of extracortical stimulation arising from afferents in the intracranial vascular system and the dura. Control of such artifactual CR's is made feasible by the fact that extracortical stimulation of this nature can generally be detected by changes in GSR. Such artifacts are abolished by bilateral trigeminal neurotomy, which denervates the pertinent region but does not interfere with the formation or retention of CR's established solely to cortical stimulation. Pupillary dilatation, a common response to brain stimulation, and a possible source of extracortical proprioceptives, may be controlled by the application of mydriatics to the eye.

Doty's results indicate that conditioned leg flexions may be established when stimulation of points in the marginal, postlateral, suprasylvian, or ectosylvian gyri serves as the CS. (These areas include the primary visual and auditory cortex, and stimulation of their homologues in man elicits visual and auditory sensations [88].) CR's once established may be maintained over periods up to three months following suspension of training. It is difficult to interpret the results on differentiation, generalization, and the differential effectiveness of cortical versus peripheral CS's because of the extreme individual variability reported. (The criterion used, 15 out of 25, seems low in view of such high variability.) Some preliminary results on the role of subcortical structures in intraneural conditioning have been reported (**93**). These findings suggest that undercutting of the stimulated areas may produce severe

deficits in established CR's. Circumsection or isolation of such areas appears to have negligible effects upon established CR's.

These results, though preliminary, indicate that electrical stimulation of the brain may be of considerable value in research on the neural mechanisms involved in the conditioned response.

Neural Mechanisms in Learning and Motivation

Perhaps the most noticeable trend in current research on brain stimulation and learning is the shift in emphasis from cortical to subcortical structures. Although the importance of these structures has long been recognized, certain recent findings in neurophysiology have been of particular interest to physiological psychologists. Studies of diencephalic and reticular "activating systems" suggest that these "systems" play an important role in the regulation of cortical activity. Stimulation of certain diencephalic and mesencephalic structures in intact animals elicits an "alert" EEG, and behavior which has been described as "apprehension" or "alertness" (**47, 70**). Lesions in these structures result in an EEG characteristic of sleep, and a striking paucity of behavior (**57, 58**). Stimulation of other structures in these regions, and in the rhinencephalon, elicits a number of highly integrated behavior patterns as well as a wide variety of complex somatic and autonomic responses (**51, 69**). For a general review of this research, the reader is referred to papers by Lindsley (**60**), Teuber (**99**), Gloor (**36**), Magoun (**70**), Jasper (**48, 50**), and Stellar (**98**), and to the UNESCO symposium on "Brain Mechanisms and Consciousness" (**16**).

On the assumption that the interaction of certain cortical and subcortical mechanisms may underlie such psychological phenomena as consciousness, emotion, motivation, attention, etc., several theorists have speculated upon the implications of these physiological findings for behavior theory (**39, 40, 59, 87**). The studies to be reviewed below are based upon hypotheses suggested by these findings.

Reinforcement Produced by Electrical Stimulation of the Brain

Negative Reinforcement

Over a period of 30 years, Hess (**41, 42, 43, 44, 45, 46**) has used brain stimulation of unanesthetized, intact animals to map the entire diencephalon, recording photographically the responses elicited, and verifying the electrode placements histologically. In addition to affecting such functions as circulation, respiration, temperature regulation, and digestion (**36, 90**), stimulation of this region may also elicit integrated behavior patterns. One of these patterns, the "affective-defense reaction" (**42**), has long been familiar to psychologists as the "sham-rage" or "pseudo-affective" reaction of the Cannon-Bard theory of emotion (**2, 3**). This pattern may be elicited by electrical stimulation of a zone centered in the perifornical region of the hypothalamus and extending rostrally to the medial preoptic area and the base of the septum (**36**). Hess has described the resulting behavior as follows:

"All the phenomena of growling, ear retraction, unsheathed claws, piloerection, tail lashing, etc; are present. At this stage of general excitement, a slight movement on the part of the observer is enough to make him the object of a brisk and well directed attack. . . . At the end of the stimulation, there is a rapid breakdown of the syndrome; even

the most impulsive attack is stopped, and the animal is, as a rule, willing to be petted. But the threshold for angry behavior is maintained at a low level for a longer period" (**45**). In discussing the differentiation which many authors have made between "normal" and "sham-rage" reactions, Hess notes:

"This differentiation does not apply to our observations, since the reactions of the animals during hypothalamic stimulation differed in no way from those observed by us under normal environmental conditions. . . . Our observations . . . are based on criteria identical with those conventionally applied to the phenomena of expressive mechanisms" (**45**).

For a time, observations of emotional behavior elicited by hypothalamic stimulation were used to support the concept that the hypothalamus was a "center" for emotion. However, as the conceptual pendulum swung from the extremes of strict localization theory to the extremes of field theory, many investigators began to question the meaningfulness and utility of the general concept of "centers." In an attempt to evaluate the behavioral status of some of the responses produced by hypothalamic stimulation, and to clarify the role of this structure in emotional behavior, several conditioning and learning studies have been carried out. The first of these experiments was performed by Masserman, in an attempt to demonstrate the inadequacy of the concept that the hypothalamus was the "center" for emotional behavior. It was Masserman's contention that hypothalamic stimulation produced only the external manifestations of anger, and not "a true affective experience, which would be anticipated at the sensory signal and for which the animal could learn to prepare, compensate, or adapt" (**73**).

After demonstrating that cats could learn an escape or avoidance response to peripheral shock, Masserman (**73**) attempted to condition emotional responses elicited by hypothalamic stimulation. A variety of CS's were used, and paired with hypothalamic stimulation. However, though the characteristic "emotionality" invariably followed such stimulation, the results indicated no escape or avoidance conditioning to the CS. Masserman interpreted these results as indicating that "direct stimulation of the hypothalamus induces dramatic mimetic-emotional effects, but these are not accompanied by the experiential or conative components of true rage or fear" (**73**).

A number of factors mitigate against the acceptance of such an interpretation. Precise control of electrode localization and stimulation parameters was lacking. In an area such as the hypothalamus, with its complex arrangement of nuclei in close proximity, minute variations in locus or stimulation parameters may result in markedly different response patterns. Indeed, Masserman's description of the behavior of his animals differs in a number of respects from the observations of Hess. Masserman reports that the behavior was maladaptive, was not specifically directed against objects in the environment, had little effect upon ongoing behavior, and showed little or no emotional aftereffects when hypothalamic stimulation was terminated. These differences suggest that Masserman may not have been dealing with the total "affective-defense reaction" as Hess has described it. It is also possible that failure to achieve successful avoidance conditioning was due to the fact

that the experimental situation did not permit the animals to establish a specific avoidance or escape response during hypothalamic stimulation. In instrumental avoidance conditioning, such a response is followed by termination of the shock. Lacking pretraining, Masserman's animals possessed no specific instrumental response.

Recent studies of avoidance conditioning to hypothalamic stimulation suggest that the latter factor may have been, in part, responsible for Masserman's negative results. Using a shuttle box, Cohen, Brown, and Brown (**14**) conditioned an escape response (jumping the barrier) to hypothalamic stimulation. Within 200 trials, this behavior was developed into a consistent avoidance response to a previously neutral CS. These investigators interpret their results in drive-reduction terms; they view hypothalamic stimulation as a drive-arousing or energizing stimulus, whose termination produces the reinforcement for learning. It is their contention that the behavior elicited when hypothalamic stimulation is used as the UCS in a learning situation would depend upon the characteristics of the experimental situation. That is, avoidance, approach, and perhaps other types of learning could be demonstrated in the same animal using identical electrode localization and stimulus parameters. By contrast with hypotheses formulated in terms of "the experiential or conative components of true rage and fear" (**73**), the Cohen and Brown hypothesis is accessible to experimental verification.

Another group of studies of instrumental avoidance learning motivated by brain stimulation is based upon the work of Delgado. This investigator has been concerned with the elucidation of the "cerebral structures involved in the transmission and elaboration of noxious stimulation" (**20**). Delgado's observations were made on chronic preparations, stimulated through electrodes in various subcortical and cortical structures. Stimulation of the tegmentum, VPL nucleus of the thalamus, amygdala, hippocampus, crus of the fornix, and periacqueductal gray elicits response patterns characteristic of pain and/or emotion. These responses include a variety of offensive and defensive behaviors, self-inspection, vocalization, and autonomic reactions; certain of them appear to form an integrated syndrome which Delgado calls "conditioned anxiety" (**20**). This syndrome is characterized by "anxiety" responses to the observation platforms on which the animals have previously been stimulated, and a "permanent state of distrust" manifested toward observers or other animals present during such stimulation. Observations such as these led to experiments designed to determine whether this type of central stimulation could be conditioned, and used to motivate escape and avoidance behavior.

It was found (**18**) that cats that had previously learned a wheel-turning avoidance response to shock and CS, would continue to perform this response when brain stimulation was substituted for the CS. This response, in turn, was itself easily conditioned to a previously neutral CS. Similarly, an emotional disturbance produced by central stimulation could be conditioned to one of two distinctive compartments, and could be used as the UCS to motivate escape from the "frightening" into the "safe" compartment. Electrical stimulation of the brain was also found to be an effective punishment in condi-

tioning hungry cats to avoid food. A recent study on the monkey (**22**) indicates that such stimulation is an effective substitute for a CS in reinforcing instrumental shock-avoidance responses previously conditioned to the CS. Monkeys were trained to avoid peripheral shock by overturning the left cup to a CS of a high tone. Turning over the right cup in response to a low tone or no tone was reinforced with food. Brain stimulation was then administered during the no-tone trials, and it was found that such stimulation elicited avoidance responses previously conditioned to high tone and shock. These responses are characterized as Conditioned Fear Responses (CFR). The authors advance the interpretation that "electrical stimulation . . . induces in the animal a condition similar to that which is present when it is anxious or afraid of being hurt" (**22**).

Such an interpretation, as the authors themselves admit, is only one of a number of possible interpretations. The most plausible of these alternate interpretations is that the stimulation was exciting pain fibers and therefore "elicited a pain-escape rather than a pain-avoidance response, (CFR)" (**22**). There is a considerable body of anatomical, physiological, and behavioral evidence which supports such an interpretation. Many of the structures stimulated are related to sensory fibers mediating pain, through their connections with the trigeminal nerve, the spinothalamic tract, and the VPL nucleus of the thalamus. The observations on protective movements, self-inspection and vocalization suggest the possibility of noxious stimulation, as does the fact that these responses are considerably reduced by anesthesia (**20**). Furthermore, the "anxiety" and "permanent state of distrust" observed by Delgado in cats and monkeys undergoing central stimulation has also been observed in rats subjected to prolonged or intensive *peripheral* shock (**78**). These animals display considerable aggression, directed toward observers, other animals, or portions of their cage. In short, there is reason to believe that in the case of many of the structures involved, electrical stimulation is equivalent to a repetition of the original, painful, UCS; we are thus dealing with instrumental escape, rather than instrumental avoidance conditioning.

Such an hypothesis will account for most of the observed results. However, stimulation of the hippocampus, the amygdala, the fornix, and the periacqueductal gray also served to reinforce learning, without concurrently eliciting responses characteristic of pain. Information on the anatomical and functional relationships of these structures is scanty. It is known that neurons of the mesencephalic root of the trigeminus border on the periacqueductal gray and might have been affected by stimulus spread. Furthermore, several investigators have recorded electrical responses in the hippocampal region, following noxious stimulation of unanesthetized animals (**68**), and "fear" responses have been reported to stimulation of the amygdala and hippocampus (**51**). The evidence from human observations is completely contradictory. According to Jasper (**49**), Penfield has never elicited anything resembling an emotional response to stimulation of the amygdala, the hippocampus or the hippocampal gyrus. Heath (**38**), on the other hand, reports the following observations on patients stimulated through electrodes implanted in these regions:

"Stimulation of the amygdaloid nucleus resulted in intense emotional

reactions which varied from one stimulation to the next, though parameters of stimulation were kept constant. Sometimes this stimulation produced rage, sometimes fear. The patient's reaction was, 'I didn't know what came over me. I felt like an animal.' Stimulation of the hippocampus has produced anxiety. . . ." Differences in procedure and problems of localization make comparison of the human results difficult.

Perhaps the best generalization that can be made in connection with negative reinforcement produced by brain stimulation is that, "The emotional disturbance capable of motivating learning is not produced indiscriminately by electrical stimulation of any area of the brain, but is limited to the stimulation of specific areas" (**18**). Further experiments are needed to clarify the behavioral status of the responses produced by stimulation of these areas. Such clarification will not be achieved if these experiments continue to be formulated in terms of the subjective aspects of the animals' experience during electrical stimulation. The fact that they have been so formulated suggests that inadequate consideration has been given to the complex conceptual problems involved in research on emotional behavior. For a detailed discussion of these problems, the reader is referred to papers by Hebb (**40**), and Lindsley (**59**).

Positive Reinforcement

The first observations on this phenomenon were made by Olds (**83**) in the course of studies on the effects of electrical stimulation of the reticular formation. By accident, electrodes implanted in one of the experimental animals landed, not in the reticular formation, but in the septal region of the rhinencephalon. When stimulated in this area, the rat kept returning to the corner of the observation box where it had last been stimulated. By appropriate manipulation of this effect, it was found that the animal could be directed to any part of the box. (This observation has since been confirmed in the monkey [**21**].) Animals with electrodes planted in this area were then tested in the Skinner box and it was found that rats who received electrical stimulation of this region following lever-pressing responses, would repeatedly make such responses. Moreover, they maintained extraordinarily high rates of responding, without additional forms of reinforcement (**79**). These findings were subsequently confirmed and extended to other species. It has since proved possible to develop and maintain stable lever-pressing rates in cats, rats, and monkeys for periods up to six months (**7, 95**).

Several investigators have explored the relationships between the reinforcing effects of "intracranial self-stimulation" and learning motivated by primary reinforcers such as food, water, or sex. Variations in reinforcement schedule (continuous, fixed ratio, variable interval) have been found to result in differential lever-pressing curves similar to those obtained with corresponding variations in food or water reinforcement schedules (**95**). Rate of responding is sensitive, within limits, to changes in the intensity of electrical stimulation ("amount of reinforcement"), and is negatively correlated with the duration of the interval between intracranial stimulations (**7**). Interactions between the reinforcing properties of intracranial self-stimulation and those of other reinforcers have also been demonstrated. Animals under food deprivation show higher response rates to intracranial self-stimulation than do animals under

satiation conditions (**7**, **8**). When both brain-stimulation and water-reinforcement levers are available on a continuous reinforcement schedule, the decrease in response rate to the water lever, as a function of satiation, is accompanied by a simultaneous decrease in rate of responding to brain stimulation. If water reinforcement is given on a variable-interval schedule, while the brain-stimulation schedule remains continuous, response rate on both bars is maintained at a relatively high level (**7**, **8**).

Interactions between emotional behavior and intracranial self-stimulation have also been reported. Using a Conditioned Emotional Response (CER) superimposed upon an established lever-pressing response, Brady and his associates have begun to explore these interactions (**5**, **6**, **7**). After a stable lever-pressing response has been developed in rats or cats to water reinforcement, the CER is produced by a series of conditioning trials in which an auditory stimulus is followed by shock. Within a short time, emotional behavior (crouching, defecation, immobility) has become conditioned to the CS; the extent of the conditioning is reflected in decreased response rates following the CS. However, stable CER's, originally established to water reinforcement, fail to appear when intracranial self-stimulation is substituted for water. Furthermore, when half-hour intervals of water reinforcement and intracranial self-stimulation are alternated within a single two-hour session, the CS does not elicit the CER during periods of brain stimulation, even though termination of the CS is accompanied on each trial by shock. It has also been found that resistance to the original development of the CER may be produced by providing the animal simultaneous access to a lever delivering brain stimulation, during the process of emotional conditioning to water reinforcement (**7**, **9**). The relation of such findings to the reported effects of subcortical lesions upon the CER (**6**, **9**) is not yet clear.

When an auditory CS is paired with intracranial self-stimulation, without shock, the lever-pressing rate begins to increase during the five-minute period following the CS, before the delivery of brain stimulation. With repeated trials, this increase becomes consistent and reliable, though response rates remain at significantly lower levels during the intervals between presentation of the CS (**7**). These findings suggest that intracranial self-stimulation may be used as a secondary as well as a primary reinforcer in such situations.

Olds (**81**) reports an experiment intended to compare the effectiveness of food versus brain stimulation in reinforcing maze and runway behavior in rats. Latency differences between food and stimulation groups were not significant in the runway. In the maze, however, differences between the number of trials to criterion were significant and favored the food group. A number of subsidiary observations are of some interest. While the performance of the food group improved on the first run of each succeeding day, the stimulation group showed a decrement. In a number of cases, electrode placements which had produced high response rates in the Skinner box failed to adequately reinforce maze and runway performance. Because of the problems involved in equating the groups with respect to the other variables in maze learning, the two groups are not strictly comparable.

Experiments of Brady and his associates indicate that the effectiveness of intracranial self-stimulation as reinforcement is related, not only

to electrode locus, but to the locus of prior brain stimulation as well. The nature of these relationships has been explored in the monkey, using multilead electrodes which permit simultaneous access to as many as 20 stimulation or recording points. It has been found that although high, stable response rates may be obtained from intracranial self-stimulation in the amygdala, prior stimulation in the medial forebrain bundle or orbital frontal region adversely affects these response rates. Similarly, prior stimulation of the hypothalamus or amygdala results in significantly lower and less stable rates of responding to intracranial self-stimulation of the medial forebrain bundle or entorhinal area (**7**).

It has become increasingly apparent that electrode locus is among the most important of the variables influencing the effectiveness of reinforcement produced by brain stimulation. There is, for example, some evidence that the loci of electrode placements effective in producing such reinforcement may differ in the rat, the cat, and the monkey (**7, 82, 95**). However, exposition of the findings on anatomical localization, species differences, and behavioral effects is complicated by the fact that studies of intracranial self-stimulation are often not comparable with respect to methodology, stimulus parameters, and behavioral criteria. On the basis of the studies reported thus far, only one anatomical generalization may be made. All the structures involved in this phenomenon may be included in or related to the so-called "limbic system." Several excellent reviews of its anatomy, physiology, and behavioral implications are available (**9, 25, 51, 67, 69, 85, 89**).

As might be expected, the discovery of positive reinforcement produced by intracranial self-stimulation has aroused considerable interest and a great deal of speculation. This speculation ranges from Delgado's concept of "attractive" and "unattractive" cerebral areas (**21**), to Olds' tentative conclusion that "stimulation in these areas must excite some of the nerve cells that would be excited by satisfaction of the basic drives—hunger, sex, thirst, and so forth" (**83**). The analysis of neural and behavioral mechanisms mediating the reinforcing effects of intracranial self-stimulation has hardly begun. Indeed, it has not yet been conclusively demonstrated that we are dealing with "pleasure centers" (**83**) or "a system within the brain whose function it is to produce a rewarding effect upon behavior" (**79**).

In the Skinnerian sense of a stimulus whose presentation increases the probability of occurrence of a prior response, intracranial self-stimulation must be termed reinforcing. However, the use of intracranial self-stimulation as reinforcement in a wider variety of learning situations would be desirable. To date, these reinforcing properties have been demonstrated only in lever-pressing and maze learning. Even in these situations, certain possibly significant differences may be noted between food and water reinforcement and intracranial self-stimulation.

1. Extinction of the lever-pressing response is extremely rapid, following the termination of intracranial reinforcement and there seems to be an absence of spontaneous recovery (**79**).

2. Rats could not learn a maze or runway if this involved running out of a starting alley across an open field into a goal box (**85**).

3. The stimulation group in the maze experiments showed a decrement on the first trial of each succeeding day, while the food group showed improved performance (**81**).

4. Electrode placements which had produced high response rates in the Skinner box failed to adequately reinforce maze learning (**81**).

5. In contrast to their performance under food reinforcement, monkeys undergoing intracranial self-stimulation did not respond satisfactorily on a schedule in which reinforcement was contingent upon the ***withholding*** of lever-pressing responses. Animals could not regularly delay their responses for the twenty-second interval required to obtain reinforcement (**10**).

These differences may have important implications for the interpretation of the positive reinforcement produced by brain stimulation. They suggest that it may be premature to conceptualize intracranial self-stimulation in terms of "approach," "reward," or "pleasure" centers in the brain. Such "subcortical phrenologizing" is no substitute for careful experimental analysis, and is no more likely to advance our knowledge of brain function than was "cortical phrenology."

Brain Stimulation and Learning: Miscellaneous Studies

The following studies are not directly concerned with intraneural conditioning or reinforcement produced by brain stimulation. They have in common only the fact that they are designed to explore the effects of brain stimulation during performance in a learning situation.

In 1934 Gengerelli (**28**) presented a neurophysiological theory of learning based primarily upon the effects of two parameters of electrical stimulation: frequency and pulse duration. Since that time, a considerable body of research on the behavioral and physiological effects of variations in stimulus parameters has been published (**54, 74**). Moreover, Gengerelli has recently developed a technique which permits remote-control stimulation of unanesthetized, freely running animals (**29, 30**). Electrical stimulation is transmitted to implanted electrodes by means of miniature receiving sets mounted directly on the animals. Employing this technique, Gengerelli has undertaken to test his theory in a number of learning situations, with rats. These experiments (**31, 32, 33**) are the first of a series designed to investigate the effects upon learning of variations in frequency, pulse duration and cerebral locus; it is hoped that eventually the interaction of these variables may also be studied.

As formulated by Gengerelli, the theory predicts only that stimulus frequency and duration will be important variables in determining the effects of electrical stimulation of the brain upon learning. The experiments to date have been concerned with the differential effectiveness of two frequency rates upon performance in the maze (**31**) and discrimination box (**32, 33**). The results are equivocal, indicating only that stimulation of the rat brain during performance may have some effect upon learning. The nature and direction of this effect is as yet undetermined. Further evaluation of the experiments is complicated by a number of factors. The possibility of nociceptive stimulation was not adequately controlled, and may account for most of the reported results. Electrode placements varied from animal to animal, and histological verification of electrode locus is lacking. No attempt was made to equate the effective stimulus intensity received by each animal, and this factor may be important in view of the finding (**30**) that both patterns of response and thresholds for their elicitation change markedly over time. The ingenuity

and potential utility of the Gengerelli technique is such that it is to be hoped that these factors will be more carefully controlled in future experiments.

Chiles (**12**) has reported an experiment on stimulation of the "diencephalic activating system" during performance in a lever-pressing task. After having reached an asymptote on this task, cats were stimulated through electrodes implanted in the VM and CM nuclei of the thalamus, and in the posterior hypothalamus. Such stimulation resulted in a decrease in lever-pressing and an increase in response variability. Chiles attributes the results to "distraction" produced by the sensory and/or motor consequences of the stimulation. Sensory effects from stimulation of these areas in humans, have been reported (**38**).

In an attempt to delineate the structures involved in the behavioral effects of frontal lobe ablations or prefrontal lobotomy, Rosvold and Delgado (**91**) stimulated different levels of the anterior portion of the monkey brain during performance on delayed alternation and visual discrimination tests. Although such stimulation produced responses ranging from hypo- to hyperactivity, there was no interference with the visual discrimination. Moreover, only those animals stimulated in the region of the head of the caudate nucleus showed any deficits in delayed alternation. The caudate group, however, was significantly different from and inferior to the noncaudate animals on this problem; their performance fell to chance and remained at this level for the duration of the electrical stimulus. Caudate stimulation resulted in a decrease in activity which the authors describe as "dozing" (**91**). Akert and Andersson (1) have observed "inertness" as a consequence of caudate stimulation of the cat, and such inertness could be produced in caudate monkeys by increasing the intensity or duration of the electrical stimulus (**92**). (In contrast to the hypoactivity produced by caudate stimulation, destruction of caudate points by electro-coagulation resulted in persistent hyperactivity and continued, though temporary, deficits in delayed alternation.) It has been suggested (**99**), that disturbance of postural components of behavior may play a major role in producing delayed-response deficits. If this hypothesis is correct, the effect of caudate stimulation—and destruction—upon delayed alternation may be due to interference with postural and/or orienting responses which serve as cues to the animal. The results of this experiment are provocative, and suggest that some of the deficits customarily attributed to ablation of frontal cortex may be more appropriately referred to mechanisms involving both cortical and subcortical structures.

Problems of Electrophysiological Methodology

Whenever an electrical stimulus is applied to the CNS, interpretation of the responses elicited presents a problem, not only for the psychologist, but for the biophysicist, the anatomist, and the physiologist. The biophysical aspects of this problem stem from the fact that neural tissue is extremely sensitive to variations in the parameters of electrical stimulation. Even slight changes in frequency, intensity, duration, or waveform may produce markedly different responses. Furthermore, as Lilly (**56**) has pointed out, while the fundamental of a given stimulus frequency may be affecting one structure, the harmonics of the frequency may concurrently be stimulating other areas. The intro-

duction of timers, relays, microswitches, etc., into the circuit complicates the problem of stimulus control still further. Indeed, until such factors are controlled, and stimulus parameters are systematically varied for a specific structure in a given species, no conclusions as to function can be drawn solely from the results of electrical stimulation.

The anatomical aspects of this problem have been clearly stated by Mountcastle (**77**):

"A pair of electrodes cast into the brain may fall into a volume of solid grey cells and activate those cells alone. In another position, it may encompass within its adequate stimulating field a large number of fibers connected with diverse regions of grey within the brain." In short, the physiological psychologist's aim of correlating structure with function is complicated by the fact that the site of the stimulating electrode is not necessarily synonymous with the site of action of the stimulus current.

The physiological problems in such research are also formidable. Reports of variations in impedance, threshold, and patterns of response to stimulation of chronic preparations, are common. Furthermore, intense or long continued electrical stimulation has an injurious effect upon neural tissue and may eventually produce lesions. Fortunately, progress has been made in controlling these injurious effects. Techniques for this purpose have recently been developed (**55**) and are currently being used in studies on intracranial self-stimulation. Their extension to other areas of stimulation research may be expected, if the favorable results reported to date are confirmed. In this connection, the relation between the behavioral effects of lesions and of electrical stimulation has never been clarified. The problem has assumed growing importance because of the tendency of many investigators (**9**, **80**, **85**) to interpret the behavioral effects of electrical stimulation in the light of effects produced by lesions in the structures involved. It has often been assumed, for example, that lesions and stimulation produce directly opposite effects upon behavior, but there is evidence (**91**) that such an assumption may represent a considerable oversimplification of the problem.

One final point is worthy of note. As it has been used to date, electrical stimulation does not constitute "physiological stimulation." That is, the characteristics of the stimulus are such that its effects are not necessarily equivalent to normal, physiological excitation of nerve fibers. Its effects may be due either to excitation, inhibition, or interference with ongoing neural processes.

Factors such as those summarized above represent only a portion of the problems involved in the use of electrical stimulation of the brain. They suggest that functional interpretations cannot be made solely on the basis of stimulation experiments. They indicate further that extreme methodological sophistication, however difficult to achieve, is a major prerequisite for competent research in this area. The growth of interdisciplinary research projects attests to the increasing recognition of this fact.

Conclusions

The psychologist employing brain stimulation as a technique for the study of the learning process is faced with several tasks. These include the experimental analysis of the behavior produced by such stimulation, the elucidation of physiological mechanisms involved in this behavior, and the solution of methodological, conceptual, and theoretical problems.

Progress in this area, then, may be discussed in terms of each of these tasks. In view of the fact that the majority of the studies reviewed in this paper are of recent origin, it is not surprising that few generalizations may as yet be made concerning the neurophysiological correlates of learning. Certain accomplishments and trends may be pointed out, however, and their implications for future research noted.

The use of brain stimulation as a behavioral research technique has opened new areas of research on the learning process. There has been a considerable accumulation of facts, and major advances have been made in electrophysiological methodology and instrumentation. Carefully controlled studies of cortical conditioning are being carried out, and the interaction of cortical and subcortical mechanisms in such conditioning is being investigated. Electrical stimulation of the brain has proved a useful technique in the delineation of cerebral structures related to emotion and nociception and the physiological correlates of motivational states are also being explored. Negative and positive "reinforcement" produced by brain stimulation have been demonstrated, and the relation of such reinforcement to a wide variety of behavioral and physiological variables is being studied.

Detailed experimental analysis of the effects of brain stimulation upon learning and motivation has scarcely begun. In addition to its effects upon the performance change being measured, e.g., response rate, such stimulation may also produce alterations in the general behavior of the animal. In the types of learning situations used to date, notably lever-pressing, such behavior tends to be overlooked completely or only casually observed. Analysis of such alterations in other aspects of the animals' behavior should prove fruitful in providing us with insights into the behavioral mechanisms mediating the effects of brain stimulation upon learning.

In view of the avowed purpose of most of the studies reviewed in this paper, it is surprising that so few of them have actually been concerned with the elucidation of the physiological mechanisms involved in learning and motivation. Except inferentially, we have hardly begun to investigate the neural and biochemical processes which intervene between the stimulation of a given structure and the performance changes which we are measuring. Some of the techniques which would make such research feasible are already in existence (**7**, **26**, **25**). These techniques permit the recording of electrical activity in diverse regions of the nervous system, and the assessment of biochemical changes produced by electrical stimulation. Such procedures, when combined with electrical stimulation of the CNS, may make it possible to trace the complex chain of physiological events involved in learning. The effects upon these physiological processes of manipulating certain behavioral variables may be similarly investigated. The research possibilities suggested by such techniques are among the most exciting yet encountered in physiological psychology.

Perhaps our greatest need, however, is for a type of conceptual clarity which can be achieved only through continual re-examination of the assumptions upon which our research is based. As Teuber has pointed out:

"We still find uncritical references to one or another mode of parcellation of cortex or subcortex, as if the parcellations themselves ('primary areas', 'association cortex', 'visceral

brain') yielded functional systems. Our greatest hazard, however, lies in the use of inadequate terms in the description of altered behavior. What are 'amnesia', 'tameness', 'savageness', 'visual discrimination deficit', except mere symptoms which await further experimental analysis" (**99**). Teuber's comments were directed to the general area of physiological psychology, but the parcellation and inadequate terminology to which he refers are particularly prevalent in current research on brain stimulation and learning: viz., "reward systems," "pleasure centers," "attractive and unattractive cerebral areas," "distraction," etc. Many psychologists working in this area have either ignored conceptual problems completely, or treated them in a superficial and inadequate manner.

Lashley (**52**), Hebb (**40**), Teuber (**99**), and others have called attention to these conceptual inadequacies and stressed the need for continual revision of the psychologist's "conceptual nervous system." Sophisticated theoretical formulations may not yet be possible, but adequate conceptualizations are a necessity. To cite Teuber once again, "No degree of refinement of ablation or stimulation techniques can substitute for clarity of concepts relating to structure and function" (**99**).

Whatever the status of past and present research in this area, there can be little doubt of the potential utility of brain stimulation as a technique for the study of the learning process. This utility will be considerably increased if brain stimulation is combined with electrical recording, ablation, and biochemical assessment techniques. The methodological problems to be solved are complex, but considerable progress has already been made and further advances may be expected. If experimental ingenuity and conceptual sophistication keep pace with progress in methodology and instrumentation, we may look forward to the eventual development of a psychophysiology of learning and motivation.

REFERENCES

1. AKERT, K., & ANDERSSON, S. Experimentelle Beiträge zur Physiologie des Nucleus Caudatus. *Acta physiol. Scand.*, 1951, **22**, 281–298.
2. BARD, P. A diencephalic mechanism for the expression of rage with special reference to the sympathetic nervous system. *Amer. J. Physiol.*, 1928, **84**, 490–515.
3. BARD, P. Central nervous mechanisms for emotional behavior patterns in animals. *Res. Publ. Ass. nerv. ment. Dis.*, 1939, **19**, 190–218.
4. BIRCH, H. G., & BITTERMAN, M. E. Reinforcement and learning: the process of sensory integration. *Psychol. Rev.*, 1949, **56**, 292–308.
5. BRADY, J. V., & HUNT, H. F. An experimental approach to the analysis of emotional behavior. *J. Psychol.*, 1955, **40**, 313–324.
6. BRADY, J. V. Emotional behavior and the nervous system. *Trans. N.Y. Acad. Sci., Ser. II*, 1956, **18**, 601–612.
7. BRADY, J. V. Motivational-emotional factors and intracranial self-stimulation. In D. E. Sheer (Ed.), *Electrical stimulation of the brain.* (*Subcortical integrative systems.*) Houston: Univer. of Houston Press, 1957.
8. BRADY, J. V., BOREN, J. J., CONRAD, D. G., & SIDMAN, M. The effect of food and water deprivation upon intracranial self-stimulation. *J. comp. physiol. Psychol.*, 1957, **50**, 134–137.
9. BRADY, J. V. The paleocortex and behavioral motivation. In H. F. Harlow &. C. N. Woolsey (Eds.), *Interdisciplinary research in the behavioral, biological, and biochemical sciences.* Madison: Univer. of Wisconsin Press, in press.
10. BRADY, J. V. Personal communication. Dep. of Psych., Walter Reed Army Inst. of Res., Walter Reed Army Medical Center, Washington, D.C.

11. BROGDEN, W. J., & GANTT, W. H. Intraneural conditioning: Cerebellar conditioned reflexes. *Arch. Neurol. Psychiat.*, 1942, **48**, 437–455.
12. CHILES, W. D. Performance during stimulation of the diencephalic activating system. *J. comp. physiol. Psychol.*, 1954, **47**, 412–415.
13. CLARK, J. W., & WARD, S. L. Electrical stimulation of the cortex cerebri of cats. Responses elicited in chronic experiments through implanted electrodes. *A.M.A. Arch. Neurol. Psychiat.*, 1937, **38**, 927–943.
14. COHEN, B. D., BROWN, G. W., & BROWN, MARJORIE L. Avoidance learning motivated by hypothalamic stimulation. *J. exp. Psychol.*, 1957, **53**, 228–233.
15. CULLER, E. A. Observations on direct cortical stimulation in the dog. *Psychol. Bull.*, 1938, **35**, 687–688.
16. DELAFRESNAYE, J. F. (Ed.). *Brain mechanisms and consciousness.* Oxford: Blackwell Scientific Publications, 1954. Pp. 556.
17. DELGADO, J. M. R. Permanent implantation of multilead electrodes in the brain. *Yale J. Biol. Med.*, 1952, **24**, 351–358.
18. DELGADO, J. M. R., ROBERTS, W., & MILLER, N. Learning motivated by electrical stimulation of the brain. *Amer. J. Physiol.*, 1954, **179**, 587–593.
19. DELGADO, J. M. R. Evaluation of permanent implantation of electrodes within the brain. *EEG clin. Neurophysiol.*, 1955, **7**, 637–644.
20. DELGADO, J. M. R. Cerebral structures involved in the transmission and elaboration of noxious stimulation. *J. Neurophysiol.*, 1955, **18**, 261–275.
21. DELGADO, J. M. R., & BURSTEIN, B. Attraction and avoidance evoked by septal and rhinencephalic stimulation in the monkey. *Fed. Proc.*, 1956, **15**, 143.
22. DELGADO, J. M. R., ROSVOLD, H. E., & LOONEY, E. Evoking conditioned fear by electrical stimulation of subcortical structures in the monkey brain. *J. comp. physiol. Psychol.*, 1956, **49**, 373–379.
23. DOTY, R. W., RUTLEDGE, L. T., & LARSEN, R. M. Conditioned reflexes established to electrical stimulation of cat cerebral cortex. *J. Neurophysiol.*, 1956, **19**, 401–415.
24. FRITSCH, G., & HITZIG, E. Ueber die elektrische Erregbarkeit des Grosshirns. *Arch. Anatom. Physiol. Lpz.*, 1870, **37**, 300–332.
25. FULTON, J. F. The limbic system: a study of the visceral brain in primates and man. *Yale J. Biol. Med.*, 1953, **26**, 107–118.
26. GALAMBOS, R., SHEATZ, G., & VERNIER, V. G. Electrophysiological correlates of a conditioned response in cats. *Science*, 1956, **123**, 376–377.
27. GANTT, H. W. Contributions to the physiology of the conditioned reflex. *A.M.A. Arch. Neurol. Psychiat.*, 1937, **37**, 848–855.
28. GENGERELLI, J. A. Brain fields and the learning process. *Psychol. Monogr.*, 1934, **45**, 1–113.
29. GENGERELLI, J. A., & KALLEJIAN, V. Remote stimulation of the brain in the intact animal. *J. Psychol.*, 1950, **29**, 263–269.
30. GENGERELLI, J. A. Patterns of response to remote stimulation of the brain of the intact animal. *J. comp. physiol. Psychol.*, 1951, **44**, 535–542.
31. GENGERELLI, J. A., & CULLEN, J. W. Studies in the neurophysiology of learning: I. Effect of brain stimulation during runs on maze performance in the white rat. *J. comp. physiol. Psychol.*, 1954, **47**, 204–209.
32. GENGERELLI, J. A., & CULLEN, J. W. Studies in the neurophysiology of learning: II. Effect of brain stimulation during black-white discrimination on learning behavior in the white rat. *J. comp. physiol. Psychol.*, 1955, **48**, 311–319.
33. GENGERELLI, J. A., & MOWER, R. D. Studies in the neurophysiology of learning: III. Further data on the effect of brain stimulation during black-white discrimination on learning behavior in th ewhite rat. *J. comp. physiol. Psychol.*, 1956, **49**, 513–515.
34. GIRDEN, E. Some neurological correlates of behavior. In H. Helson (Ed.), *Theoretical foundations of psychology.* New York: van Nostrand, 1951.
35. GIURGEA, C. Die Dynamik die Ausarbeitung einer zeilchen ¡Beziehung durch direkte Reizung der Hirnrinde. *Ber. ges. Physiol. exp. Pharmakol.*, 1955, **175**, 80. (Abstract)
36. GLOOR, P. Autonomic functions of the diencephalon (a summary of the experimental work of Prof. W. R. Hess). *A.M.A. Arch. Neurol. Psychiat.*, 1954, **71**, 773–790.
37. HEATH, R. G. (Ed.). *Studies in schizophrenia: A multidisciplinary approach*

to mind-brain relationships. Cambridge, Mass.: published for the Commonwealth Fund by Harvard Univer. Press, 1954.

38. HEATH, R. G. Correlations between levels of psychological awareness and physiological activity in the central nervous system. *Psychosomat. Med.*, 1955, **17**, 383–395.
39. HEBB, D. O. *The organization of behavior.* New York: Wiley, 1949.
40. HEBB, D. O. Drives and the C.N.S. (conceptual nervous system). *Psychol. Rev.*, 1955, **62**, 243–254.
41. HESS, W. R. *Beiträge zur Physiologie des Hirnstammes: I. Die Methodik der lokalisierten Reizung und Ausschaltung subkortikaler Hirnabschnitte.* Leipzig: George Thieme Verlag, 1932.
42. HESS, W. R., & BRÜGGER, M. Das subkortikale Zentrum der affektiven Abwehrreaktion. *Helvet. Physiol. Pharmacol. Acta*, 1943, **1**, 33–52.
43. HESS, W. R. *Die functionelle Organisation des vegetativen Nervensystem.* Basel: Benno Schwabe, 1948.
44. HESS, W. R. *Diencephalon.* New York: Grune & Stratton, 1954.
45. HESS, W. R., & AKERT, K. Experimental data on the role of the hypothalamus in the mechanism of emotional behavior. *A.M.A. Arch. Neurol. Psychiat.*, 1955, **73**, 127–129.
46. HESS, W. R. *Hypothalamus and thalamus: Documentary pictures* (English and German text). Stuttgart: George Thieme Verlag, 1956.
47. INGRAM, W. R., KNOTT, J. R., WHEATLEY, M.D., & SUMMERS, T. D. Physiological relationships between hypothalamus and cerebral cortex. *EEG clin. Neurophysiol.*, 1951, **3**, 37–58.
48. JASPER, H. H. Diffuse projection systems: the integrative action of thalamic diffuse projection systems. *EEG clin. Neurophysiol.*, 1949, **1**, 391–419.
49. JASPER, H. H. Discussion of a paper by J. Olds. In H. F. Harlow & C. N. Woolsey (Eds.), *Interdisciplinary research in the behavioral, biological, and biochemical sciences.* Madison: Univer. of Wisconsin Press, in press.
50. JASPER, H. H. Reticular-cortical systems and theories of the integrative action of the brain. In H. F. Harlow & C. N. Woolsey (Eds.), *Interdisciplinary research in the behavioral, biological, and biochemical sciences.* Madison: Univer. of Wisconsin Press, in press.
51. KAADA, B. R. Somato-motor, autonomic and electrocorticographic response to electrical stimulation of rhinencephalic and other structures in primates, cats, and dogs. *Acta physiol. Scand.*, 1951, **24**, 1–285 (Suppl. 83).
52. LASHLEY, K. S. Functional interpretation of anatomic patterns. *Res. Publ. Ass. Res. nerv. ment. Dis.*, 1952, **30**, 529–547.
53. LILLY, J. C., HUGHES, J. R., & GALKIN, THELMA W. Some evidence for gradients of motor function in the whole cerebral cortex of the unanaesthetised monkey. *Proc. XX int. physiol. Cong.*, 1956, 567–568.
54. LILLY, J. C., AUSTIN, G. M., & CHAMBERS, W. W. Threshold movements produced by excitation of cerebral cortex and efferent fibres with some parametric regions of rectangular current pulses (cats and monkeys). *J. Neurophysiol.*, 1952, **15**, 319–342.
55. LILLY, J. C., HUGHES, J. R., ALVORD, E. C., JR., & GALKIN, THELMA W. Brief, non-injurious electric waveform for stimulation of the brain. *Science*, 1955, **121**, 468–469.
56. LILLY, J. C. Discussion of a paper by J. Olds. In H. F. Harlow & C. N. Woolsey (Eds.), *Interdisciplinary research in the behavioral, biological, and biochemical sciences.* Madison: Univer. of Wisconsin Press, in press.
57. LINDSLEY, D. B., BOWDEN, J. W., & MAGOUN, H. W. Effect upon EEG of acute injury to the brain stem activating system. *EEG clin. Neurophysiol.*, 1949, **1**, 475–486.
58. LINDSLEY, D. B., SCHREINER, L. H., KNOWLES, W. B., & MAGOUN, H. W. Behavioral and EEG changes following chronic brain stem lesions in the cat. *EEG. clin. Neurophysiol.*, 1950, **2**, 483–498.
59. LINDSLEY, D. B. Emotion. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951.
60. LINDSLEY, D. B. Physiological psychology. *Annu. Rev. Psychol.*, 1956, **7**, 323–348.
61. LOUCKS, R. B. Preliminary report of a technique for stimulation or destruction of tissues beneath the integument and the establishment of conditioned reactions with faradization of the cerebral cortex. *J. comp. Psychol.*, 1933, **16**, 439–444.
62. LOUCKS, R. B. The experimental delimitation of neural structures essential for learning: the attempt to condition striped muscle responses with

faradization of the sigmoid gyri. *J. Psychol.*, 1935, **1**, 5–44.

63. LOUCKS, R. B. Studies of neural structures essential for learning. II. The conditioning of salivary and striped muscle responses to faradization of cortical sensory elements, and the action of sleep upon such mechanisms. *J. comp. Psychol.*, 1938, **25**, 315–332.
64. LOUCKS, R. B., & GANTT, W. H. The conditioning of striped muscle responses based upon faradic stimulation of dorsal roots and dorsal columns of the spinal cord. *J. comp. Psychol.*, 1938, **25**, 415–426.
65. LOUCKS, R. B. Methods of isolating stimulation effects with implanted barriers. In D. E. Sheer (Ed.), *Electrical stimulation of the brain. (Subcortical integrative systems)*. Houston: Univer. of Houston Press, 1957.
66. LOWENBACH, H., & GANTT, H. W. Conditioned vestibular reactions. *J. Neurophysiol.*, 1940, **3**, 43–48.
67. MACLEAN, P. D. Psychosomatic disease and the "visceral brain." *Psychosomat. Med.*, 1950, **11**, 19–49.
68. MACLEAN, P. D., HOROWITZ, N. H., & ROBINSON, F. Olfactory-like responses in the pyriform area to non-olfactory stimulation. *Yale J. Biol. Med.*, 1952, **25**, 159–172.
69. MACLEAN, P. D. The limbic system ("visceral brain") in relation to central grey and reticulum of the brain stem: Evidence of interdependence in emotional processes. *Psychosomat. Med.*, 1955, **17**, 355–366.
70. MAGOUN, H. W. An ascending reticular activating system in the brain stem. *A.M.A. Arch. Neurol. Psychiat.*, 1952, **67**, 145–154.
71. MAIER, N. R. F., & SCHNEIRLA, T. C. Mechanisms in conditioning. *Psychol. Rev.*, 1942, **49**, 117–133.
72. MALIS, L. I., PRIBRAM, K. H., & KRUGER, L. Action potentials in motor cortex evoked by peripheral nerve stimulation. *J. Neurophysiol.*, 1953, **16**, 161–167.
73. MASSERMAN, J. H. Is the hypothalamus a center for emotion? *Psychosomat. Med.*, 1941, **3**, 3–25.
74. MIHAILOVIC, L., & DELGADO, J. M. R. Electrical stimulation of monkey brain with various frequencies and pulse durations. *J. Neurophysiol.*, 1956, **19**, 21–36.
75. MONNIER, M., & LAUE, H. Technique de dérivation des activités électriques corticales et sous-corticales pendant la stimulation du diencephale chez le lapin. *Helvet. physiol. Pharmacol. Acta*, 1953, **11**, 73–80.
76. MORGAN, C. T. The psychophysiology of learning. In S. S. Stevens (Ed.) *Handbook of experimental psychology*. New York: Wiley, 1951.
77. MOUNTCASTLE, V. B. Discussion of a paper by J. Olds. In H. F. Harlow & C. N. Woolsey (Eds.), *Interdisciplinary research in the behavioral, biological, and biochemical sciences*. Madison: Univer. of Wisconsin Press, in press.
78. O'KELLEY, L. I., & STECKLES, L. C. A note on long enduring emotional responses in the rat. *J. Psychol.*, 1939, **8**, 125–131.
79. OLDS, J., & MILNER, P. Positive reinforcement produced by electrical stimulation of septal area and other regions of the rat brain. *J. comp. physiol. Psychol.*, 1954, **47**, 419–427.
80. OLDS, J. Physiological mechanisms of reward. In M. R. Jones (Ed.), *Nebraska symposium on motivation*, 1955.
81. OLDS, J. Runway and maze behavior controlled by basomedial forebrain stimulation. *J. comp. physiol. Psychol.*, 1956, **49**, 507–512.
82. OLDS, J. A preliminary mapping of electrical reinforcing effects in the rat brain. *J. comp. physiol. Psychol.*, 1956, **49**, 281–285.
83. OLDS, J. Pleasure centers in the brain. *Sci. Amer.*, 1956, **195**, 105–116.
84. OLDS, J., KILLAM, K. F., & BACH-Y-RITA, P. Self-stimulation of the brain used as a screening method for tranquilizing drugs. *Science*, 1956, **124**, 255–256.
85. OLDS, J. Adaptive functions of paleocortex and related structures. In H. F. Harlow & C. N. Woolsey (Eds.), *Interdisciplinary research in the behavioral, biological, and biochemical sciences*. Madison: Univer. of Wisconsin Press, in press.
86. OLDS, J. Personal communication. Dep. of Anat. & Physiol., Univer. of Calif., Los Angeles, Calif.
87. PAPEZ, J. W. A proposed mechanism of emotion. *A.M.A. Arch. Neurol. Psychiat.*, 1937, **38**, 725–743.
88. PENFIELD, W., & RASMUSSEN, T. *The cerebral cortex of man*. New York: Macmillan, 1950.
89. PRIBRAM, K. H., & KRUGER, L. Functions of the "olfactory brain." *Ann. N. Y. Acad. Sci.*, 1954, **58**, 109–138.
90. RANSON, S. W., & MAGOUN, H. W. The

hypothalamus. *Ergebn. Physiol.*, 1939, **45**, 56–163.

91. ROSVOLD, H. E., & DELGADO, J. M. R. The effect on delayed-alternation test performance of stimulating or destroying electrically structures within the frontal lobes of the monkey's brain. *J. comp. physiol. Psychol.*, 1956, **49**, 365–372.
92. ROSVOLD, H. E. Personal communication. Section on Animal Behavior, National Inst. of Ment. Hlth., Bethesda, Md.
93. RUTLEDGE, L. T., JR., & DOTY, R. W. Neural pathways of conditioned reflexes to cortical stimulation. *Fed. Proc.*, 1956, **15**, 158.
94. SHEER, D. E. (Ed.). *Electrical stimulation of the brain. (Subcortical integrative systems.)* Houston: Univer. of Houston Press, 1957.
95. SIDMAN, M., BRADY, J. V., BOREN, J. J., CONRAD, D. G., & SCHULMAN, A. Reward schedules and behavior maintained by intracranial self-stimulation. *Science*, 1955, **122**, 830–831.
96. SNIDER, R. S., & STOWELL, A. Receiving areas of the tactile, auditory, and visual areas in the cerebellum. *J. Neurophysiol.*, 1944, **7**, 331–358.
97. SPENCE, K. W. Theoretical interpretations of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.
98. STELLAR, E. Physiological psychology. *Annu. Rev. Psychol.*, 1957, **8**, 415–436.
99. TEUBER, H.-L. Physiological psychology. *Annu. Rev. Psychol.*, 1955, **6**, 267–296.
100. WALKER, A. E. Afferent connections. In P. C. Bucy (Ed.), *The precentral motor cortex*. Urbana: Univer. of Ill. Press, 1944.
101. WOOLSEY, C. N. Patterns of representation in the cerebral cortex. *Fed. Proc.*, 1947, **6**, 438–441.

Received April 2, 1957.

PSYCHOLOGICAL BULLETIN
Vol. 54, No. 5, 1957

LABORATORY STUDIES OF BEHAVIOR WITHOUT AWARENESS

JOE K. ADAMS
VA Hospital, Palo Alto[1]

The term "behavior without awareness" has been used to cover a considerable variety of psychological processes. This review will attempt a classification of these processes in terms of the experimental conditions under which they have been observed. Many alternative classifications could be made; the chief virtue of the present one is that it tends to separate those results which can be relatively easily reproduced from those which can be reproduced only with considerable difficulty or not at all.

The following kinds of observations will not be included:

1. Studies of posthypnotic suggestion.
2. Studies of automatic writing.
3. Studies of extrasensory perception.
4. Nonexperimental observations.
5. Studies of processes of which *S*s are ***never*** directly aware.

Various aspects of behavior without awareness were reviewed in 1917 by Coover (**9**), in 1940 by Collier (**8**), in 1942 by Miller (**49**), and in 1951 by Lazarus and McCleary (**40**). The present review, however, presents a coverage somewhat different from any of these.

By *behavior* the reviewer means anything that the individual does that is publicly observable, whether the behavior is defined in terms of movement, accomplishment, or physiological changes.[2]

By *awareness* the reviewer means *conscious awareness*, which is taken as a primitive (undefined) term. Some psychologists consider the indices or criteria of awareness, e.g., appropriate verbalization, as the only meaning that can legitimately be given to the term; others use "awareness" as a construct based upon such indices but without the substantive meaning of consciousness. The clarity gained by avoiding the phenomenological meaning is more apparent than real, as no psychologist to date has given an adequate behavioral definition of awareness except in terms of rather specific experimental conditions, in other words, the very same terms in which any psychologist specifies his indices or criteria of awareness.

Lack of awareness by *S* may be with respect to:

1. The behavior itself; e.g., *S* may be whispering without realizing that he is whispering, or he may be saying "vertical" on 60% of a sequence of trials without being aware that he is responding more than 50% of the time in this way.[3]
2. The relation of his behavior to some contingent event; e.g., *S* may be unaware that his behavior is being influenced by *E*'s saying "right."
3. The sensory experiences that usually accompany a given kind of stimulation; e.g., *S* may be unaware of a very faint light that is being presented in the sense that he lacks the

[1] This manuscript was prepared while the author held the Thomas Welton Stanford Fellowship in Psychical Research at Stanford University, 1955–56.

[2] Some psychologists object to the inclusion of physiological changes under the term "behavior." There seems to be no generally accepted word which has the meaning desired.

[3] Research which deals with processes of which *S*s are ***never*** directly aware, such as change in pupillary size and GSR, will not be included unless some other kind of lack of awareness is involved.

usual visual sensations. This is one of the traditional meanings of "lack of conscious awareness."

4. The fact that he is discriminating the presence from the absence of a given stimulus better than chance (this kind of lack of awareness often follows from 3, under suitable laboratory conditions).

5. The fact that he is responding differentially to different stimuli in a particular systematic way.

6. Contingencies in the environment which produce changes in *S*'s behavior; e.g., *S* may be unaware that a given word has preceded electric shock, though the word produces changes in heart rate or GSR.

There is probably no psychologist who doubts the existence of any of the foregoing kinds of behavior without awareness, if he is allowed to express them in the terminology he prefers. For example, among clinical psychologists it is widely believed that in ordinary social interaction people often respond, in ways of which they are unaware, to stimuli which they are unaware that they are responding to, and that such behavior without awareness can be of special importance in psychopathology. In therapy great stress is placed on the importance of becoming aware of the stimuli to which one responds, of the response one makes, and of the reasons that one responds in the way one does.

Experimental psychologists demonstrate their belief in the occurrence of behavior without awareness by their insistence upon certain kinds of methodological precautions. If it is asserted that a functional relationship has been demonstrated by a given experiment, and if through experimental error or faulty design some other stimulus variable has been correlated in time with the experimental stimulus variable, the alternative hypothesis that the second stimulus may have served as a cue is usually considered quite plausible, even if the *S*s protest that they were not aware of the presence of the uncontrolled variable. For example, behavior without awareness is considered a very plausible explanation of above-chance results of those ESP experiments in which correlated variables could conceivably have served as cues (**9, 34, 35, 52**).

The impression is apparently widespread that behavior without awareness has been adequately demonstrated in the laboratory, although certain forms of it, e.g., subception (**4, 12, 29, 30, 39, 40, 44, 45, 46**) and conditioning to subliminal stimuli (**1, 2, 26, 50, 62, 74**), are generally known to be very elusive, to say the least. With regard to concept formation, for example, Leeper (**41**, p. 731) says, "Some of the *S*s, . . . , develop the ability to name new examples without being able to say how they do it, even when the necessary formulations lie well within the limits of their vocabularies . . . . In the studies of Heidbreder (1924, 1946, 1947), Smoke (1932), Reese and Israel (1935), Snygg (1935), Heidbreder, Bensley, and Ivy (1948), and Bouthilet (1948) this unconscious process is abundantly demonstrated." After discussing additional experiments, many of which involve animals, Leeper concludes, " . . . these experiments emphasize that the controlling processes involved ought to be studied primarily in terms of their functional properties and only incidentally, if at all, in terms of whether they are conscious processes" (**41**, p. 734).

As this review will show, however, only one form of behavior without awareness has been established on an unquestionable experimental basis. Experiments which reputedly have

shown the kinds of phenomena which are ordinarily meant by "behavior without awareness" have one or more of the following limitations:

1. Alternative explanations are plausible.
2. Negative results have been obtained upon replication by other experimenters.
3. No replication has been reported.
4. The effects are so slight that the question of whether certain statistical assumptions are precisely satisfied becomes critical (this is true even when by virtue of large numbers of observations the p value is very close to zero).

The latter part of the quotation from Leeper expresses an attitude that is quite prevalent, even among some experimenters who have worked on learning without awareness. It is an attitude that is directly contrary to the importance attached to awareness by those concerned with psychodynamics.

Distinctions which seem to be of special importance in a discussion of behavior without awareness are whether stimulus discrimination is involved, whether the behavior is learned during the experimental procedure or not, and whether, in case stimulus discrimination is involved, the *S* knows what it is he is supposed to be discriminating. These distinctions are utilized in the outline which follows.

I. Behavior Involving Stimulus Discrimination

A. *Experiments in Which the Subject Knows the Specific Nature of the Cue, but Is Not Aware of the Usual Sensory Qualities nor of the Fact That He Is Succeeding in Discriminating*

1. *Behavior Not Learned During the Experiment.*

The only kind of behavior without awareness which can be easily replicated is the kind reported in 1884 by Peirce and Jastrow (**52**). Working with lifted weights, they provided an example of the now familiar fact that if *S* is forced to make a discriminatory judgment, he may do far better than chance even though his judgments are made with no confidence whatsoever, e.g., he thinks that he is merely guessing because the usual kind of sensory experience is lacking. The procedure Peirce and Jastrow used was rather unusual. A weight was lowered upon the tip of *S*'s forefinger (Peirce) or middle finger (Jastrow). *S* then said, "Change," whereupon the weight was either increased or diminished by a small amount. *S* again said, "Change," whereupon the weight was restored to its original intensity. *S* judged the order of the changes and gave a confidence judgment of 0, 1, 2, or 3, 0 meaning that there was "absence of any preference for one answer over its opposite, so that it seemed nonsensical to answer." From the published report the percentages correct for judgments given with 0 confidence can be computed only for Jastrow, and are reproduced in Table 1.

In Table 1 all percentages listed under 0 confidence are significantly greatly than expectation of 50, and some are quite appreciably greater. These results are given in some detail because, although based upon only one subject, and a very sophisticated one at that, they are similar to results that have been reported by investigators using other experimental procedures with lifted weights or working in other sense modalities. For example, Fullerton and Cattell (**14**), following the more usual procedure of having *S* lift one weight and then a second and then judge whether the second was lighter or heavier than

TABLE 1

PERCENTAGE CORRECT FOR JUDGMENTS GIVEN WITH ZERO CONFIDENCE BY JASTROW, IN AN EXPERIMENT WITH TWO SETS OF LIFTED WEIGHTS

Ratio of Pressures	Confidence				Ratio of Pressures	Confidence			
	0	1	2	3		0	1	2	3
1.015	63	75	60	100	1.005	59	66	—	—
	(176)	(68)	(5)	(1)		(497)	(3)	(0)	(0)
1.030	75	87	96	100	1.010	66	52	—	—
	(141)	(83)	(24)	(2)		(558)	(62)	(0)	(0)
1.060	91	99	96	100	1.020	75	92	—	—
	(94)	(76)	(56)	(24)		(526)	(74)	(0)	(0)

Note.—The number of judgments is given in parentheses. Percentages for other degrees of confidence are given for comparison. Computed from Peirce and Jastrow, 1884.

the first, found that, of those judgments given with confidence judgment "doubtful" (which in this experiment was synonymous with "pure guess"), 65% (based on 285 judgments) were correct for one *S* and 60% (based on 405 judgments) for the other. Sidis (**60**) had *S* guess what letter or numeral was printed on a card held so far away that *S* thought he was making pure guesses; 20 *S*s obtained about 67% correct judgments of whether a letter or a number was being shown (expectation = 50%), about 25% correct judgments of the particular character (expectation = 2.8%, if one considers each of the 36 possibilities as equally probable), and about 35% correct judgments of the particular character when *S* knew which 10 characters were on the cards (expectation = 10%). Stroh, Shaw, and Washburn (**65**) replicated this experiment with 13 *S*s and obtained from 16% to 74% correct judgments (expectation = 10%). Even when the characters (letters only) were enclosed in rectangles so as to attenuate the effect of gross differences in shape or bulk, as between "B" and "T," the percentage results for 8 *S*s were 8, 18, 21, 24, 34, 37, 63, 67. These investigators also reported significant results (from 10% to 39% correct, for 10 *S*s) when the name of the letter was whispered so faintly that "no sound whatever could be heard at the distance at which the observer sat."[4]

Coover (**9**) used tachistoscopic exposures of letters or numerals; *S* indicated after each exposure whether the letter he wrote down was a perception, an inference based on partial perception, or a pure guess. In one experiment, the percentage right for pure guesses of 35 *S*s varied from 0% to 56%, with an over-all result of 9%. In a second experiment, the results for 15 *S*s varied from 0% to 14% with an over-all result of 8%. The expectation which Coover quotes in both experiments is 2.8% (one out of 36).[5] In two other experiments *S* was told the 10 possibilities. In the first, the results for 15 *S*s ran from

[4] The mean percentages are 37%, 34%, and 20%; however, the number of trials varied considerably from *S* to *S* in each of three experiments.

[5] The character exposed was always one of the following ten: B H K U Z 2 4 5 7 9. The expectation for any group of *S*s cannot be determined without a study of their guessing habits; therefore, the figure of 2.8% which Coover gives has an indeterminate error. Coover was aware of the influence of guessing habits upon expectation, and he analyzed his results to show that guessing habits could have raised the expectation to a maximum of about 4.6%, leaving the results statistically significant.

6% to 54% (over-all 13%). In the second experiment, results for 9 *Ss* were from 8% to 27% (over-all 15%). The number of trials was sufficient to be able to assert that many of the individual differences were genuine, with about 60% of the *Ss* showing the effect with expectation of 2.8% and about 20% with expectation of 10%.[6] Coover also obtained significant results with playing cards held at a distance. With 7 *Ss* the over-all percentage right for guessed color was 76% (expectation = 50%) and for number was 45% (expectation = 10%). When the names of the characters were whispered at 25 meters distance and *S* guessed, the over-all results for 4 *Ss* were 7% (2.8 < expectation < 10%) and when *S* was informed of the 10 characters, 12%. Both these results must be considered of dubious statistical significance.

Pillai (**55**) presented letters at a distance (40 *Ss*) and also by whispering (60 *Ss*) and reported that *all Ss* achieved scores significantly greater than expected by chance; the results have not been published in detail. Baker (**2**) reported as high as 81% discrimination of two lines as either vertical and horizontal or diagonal, when the 4 *Ss* reported zero confidence. In the same paper, Baker reported as high as 86% discrimination of dot-dash or dash-dot, given with an audiometer to 4 other *Ss*. In both these experiments Baker found the percentage right decreasing as the intensity was decreased below the highest point at which all confidence judgments were zero.[7]

[6] The difference can probably be accounted for by the fact that the expectation was actually higher than 2.8%, as indicated in footnote number 5.

[7] In a replication of this experiment by Hilgard, Miller, and Ohlson (**26**), no such discrepancies were obtained. See the comment in Section I, A, 2.

Williams (**76**) projected a circle, triangle, or square very faintly on a screen; *S* guessed which of the three figures was being presented; results for 11 *Ss* ranged from 15% to 70% right, with over-all 50% (expectation = 33.3%) even when they "saw nothing at all."[8] The 4 highest scoring *Ss* were rerun; the results were 45%, 45%, 50%, and 56% (over-all 49%). A control series without projection of figures gave 29%, 31%, 31%, and 33%.

Miller (**47**) used a similar procedure, with 5 naive and 5 sophisticated *Ss*. In this section we shall consider only the sophisticated *Ss*, i.e., those who knew that one of five possible geometric figures was being projected but thought nonetheless that their guesses were unaided by perception. There were 125 trials at each of 4 degrees of illumination, 60 v, 70 v, 80 v, and 90 v, all below the conscious threshold. The over-all percentages were 21, 23, 27, and 35 respectively (expectation = 20%). The highest percentage right for any *S* was 56, at 90 v. Vinacke replicated this experiment with *Ss* who were dark adapted, but obtained negative results (**73**).

Since 1939 there seems to have been little interest in this kind of discrepancy between performance and expectation, perhaps because of a declining interest among experimentalists in whether a process is conscious or not; indeed, the term "threshold"

[8] The same random series of 54 figures was exposed more than once to each *S*, only the pure guesses being analyzed; therefore, because of sequential learning, expectation may be slightly greater than 33%. This objection can also be raised to the second part of Williams' experiment and is not overruled by his control series in which no figures at all were projected (and the results were in accordance with expectation), for *S* then had no reference points for any sequential learning that may have occurred. It seems very doubtful, however, that sequential learning can account for the large results that Williams obtained.

is often defined entirely in terms of discrimination, so that the expression "subliminal discrimination," which originally meant discrimination when *S* was not consciously aware of the usual sensory qualities (and therefore, unless very sophisticated about this phenomenon, was unaware that he was able to discriminate), now seems self-contradictory, although of course it can be "translated" into a statement about co-occurrence of two kinds of behaviors, e.g., saying "I'm purely guessing" and at the same time discriminating above expectation.

Behavior without awareness of the kind described in this section can easily be obtained even as a class exercise. Nevertheless, caution must be exercised to avoid alternative explanations in terms of sequential learning (which, to be sure, might involve another type of behavior without awareness) or by inadequate randomization of the sequences of stimuli. Also, the classification "pure guess" should not be used too lightly. It is important not only to instruct *S* appropriately but to try to understand what he means when he uses this, or some similar, term. Urban (**70**) comments on this problem.

There seem to have been no published studies of learning to judge one's uncertainties more accurately (i.e., to eliminate the discrepancy between confidence judgment and performance) or to manipulate the discrepancies through practice, knowledge of results, reward, or punishment. There has also been little or no attempt to determine the average magnitudes of discrepancies between expectation and performance for various experimental conditions, or to relate these discrepancies to other variables, or to explore the relation of individual differences (which tend to be quite large) in these discrepancies to other variables.[9]

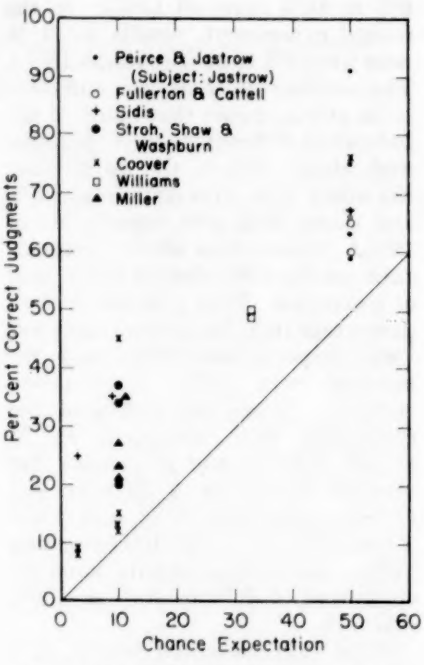

FIG. 1. DISCRIMINATION WITHOUT AWARENESS IN EXPERIMENTS IN WHICH *S* KNOWS THE SPECIFIC NATURE OF THE CUE

Figure 1 indicates some of the mean results obtained under the very diverse conditions used by the investigators whose work has been cited.

2. *Response Based on Learning During the Experiment.*

There have been relatively few laboratory studies of *learning* to respond to subliminal stimuli when *S* is aware of the specific nature of the cue. Newhall and Sears (**50**) paired very faint visual stimuli with a shock to the right hand. *S* was instructed to maintain fixation, to watch carefully for a

[9] Pauline Austin Adams and the author are at present engaged in studies of several of these problems.

visual stimulus, to report perception of the stimulus in all cases of higher certitude than a guess, and to "let the right hand take care of itself." On the test trials (without shock), there were 8 instances of conditioned finger retraction without *S*'s reporting having seen a light on the same trial. The authors believe that the most plausible interpretation was that the motor response was on these occasions more sensitive than the conscious, but they point out alternative explanations, e.g., that *S*'s attention had slipped from the visual field to the hand-shock situation and that a faint light might have passed unperceived because of lowered intensity due to lowered attention. The latter interpretation would involve behavior without awareness but would not meet the authors' criteria for subliminal conditioning. The authors are cautious in drawing any conclusion; at best the phenomenon was present only in a very weak and unstable form. Baker (**2**), after some previous unsuccessful attempts (**62**), reported very large and stable effects of pupillary conditioning to subliminal auditory stimuli, but Wedell, Taylor, and Skolnick (**74**) and Hilgard, Miller, and Ohlson (**26**) made very thorough attempts to reproduce Baker's results without success.

Lazarus and McCleary (**40**) presented 10 nonsense syllables to 9 *S*s, pairing shock with 5 of the syllables. All 10 syllables were then presented tachistoscopically at near threshold duration; the critical syllables gave significantly more GSR, even when *S* failed to identify them verbally, than did the neutral syllables. Quite aside from the question of adequate replication of this experiment and from the controversy that has arisen over the theoretical interpretation of the results (**29, 30, 39, 46**), Eriksen (**12**) has pointed out that it is possible that if the verbal response by *S* had not been restricted to one of the 10 syllables, e.g., if *S* had been allowed to say "I see a 'v' and then a blur and then a 'k' . . . ," the "subception effect" as it had been called might have been greatly reduced, perhaps to zero.

B. *Experiments in Which the Subject Believes That a Cue Is Being Given and Knows the General Nature of the Cue, e.g., That It Is Visual, but Does Not Know Its Specific Nature*

Experiments in concept formation are sometimes cited, as by Leeper in the previous quotation, as showing discrimination without awareness. In these experiments the stimulus is clearly supraliminal and *S* knows that some cue is being given and knows (at least for some of the concepts) what the general nature of the cue is. Evidence for discrimination without awareness consists of a greater number of correct identifications of instances of a concept than can be accounted for by the *S*'s verbalization (or drawing), assuming that the necessary formulation would lie within the limits of the *S*'s readily available vocabulary (or ability to draw). Actually, none of the published reports of concept-formation experiments offer proof of discrimination without awareness in this sense, although it is asserted by some of the authors to have occurred. In Hull's classic study (**31**), in which Chinese characters containing the same radical were paired with the same nonsense syllable, there is no evidence presented that the presumably inadequate drawings of radicals by the *S*s were not a sufficient basis on which to differentiate the different groups of characters. Similarly, in Smoke's study (**63**), it is not shown that the verbalizations, however much they differed from

those with which the experimenter started, would not have served to differentiate the groups of geometric drawings. In Heidbreder's 1924 study (**20**), cited by Leeper, no assertion is made that the *S*s were unaware of the criteria for making the figures (which for purposes of this discussion can be considered functionally equivalent to identifying members of a class); Heidbreder asserts merely that the *S*s were unaware of certain kinds of behaviors ("spectator" and "participant" behaviors) in which they engaged while arriving at the solutions. In Heidbreder's later papers (**21**, **22**, **23**, **24**) it is reported that correct identifications preceded correct formulations in some cases, but, as with Smoke's report, the evidence that the formulations were an inadequate basis for classification is not presented.

The Rees and Israel study (**58**), also cited by Leeper, deals with a kind of phenomenon that is at least superficially different from the kind of discrimination without awareness reported obtained in the concept-formation studies, and will be discussed in Section I, D. Of the other papers cited by Leeper, those by Snygg (**64**) and by Bouthilet, the second is unpublished and the first is unrelated to discrimination without awareness, as far as this reviewer can determine.

Goodnow and Postman (**15**) had *S*s decide which of two variations (additive and subtractive) matched a given geometric design. In the instructions the additive and subtractive variations were explained to *S*, and he was "led to believe that the correct choice depended on the nature of the designs and that there was a general principle governing correct choices" (**15**, p. 17). *E* said "Right" or "Wrong" after each choice. After every other block of 10, *S* was asked, "How did you decide which card to choose?" Six groups were run with varying proportions (.5, .6, .7, .8, .9, 1.0) of additive designs designated "Right" by *E*. The groups tended to match, in their choices, the proportions of additive designs designated "Right" by *E*, yet, "Their verbal responses gave no indication that they perceived the problem as one of probability discrimination" (**15**, p. 21). However, the *S*s were apparently not asked whether they did in fact choose additive designs more often. In view of the instructions, it is not surprising that "only three of the 40 *S*s in the partial reinforcement groups abandoned the search for a lawful solution . . . and treated the task as one of discriminating probabilities" (**15**, p. 19). As there are many hypotheses (e.g., sequential hypotheses) which produce a given proportion of additive responses (e.g., the hypothesis SAASA gives 60% additive choices) there is no evidence presented in the report that there was any discrepancy between what the *S*s did and what they said they were doing. That *S*s sometimes match probabilities by hypotheses about sequences is shown in a paper by Hake and Hyman (**18**).

C. *Experiments in Which the Subject Believes That a Cue Is Being Given, but Does Not Know Even Its General Nature*

Postman and Jarrett (**56**), using a procedure similar to that of Thorndike and Rock (see next section), had *S*s respond to each word given by *E*, saying "Right" when *S*'s word was connected sequentially with *E*'s word and "Wrong" when *S*'s word was unrelated, connected denotatively, *etc.* One group was told the principle at the end of the first block of 20 words; the uninformed group was required

to attempt a statement of the principle at the end of each block. The authors reported that those uninformed *Ss* who were able at some point during the experiment to verbalize the principle correctly showed a small but significant amount of learning prior to verbalization; the improvement was from about 2.3 to about 3.5 words right per block (about 11% to 18%). The authors do not deal with the problem of awareness of correlated hypotheses; they admit however, "the results . . . show only little reliable evidence for learning without awareness" (**56**, pp. 254–255).

Philbrick and Postman (**54**) presented a sequence of 216 words to each of 48 *Ss*, the *S* being told that a number from 1 to 9 had been paired with each word and that he was to try to guess what the number was on each trial, *E* saying "Right" or "Wrong." Most *Ss* apparently assumed that some cue was being given, especially as each *S* was asked to state the principle on which he was basing his responses after each block in which he had 4 or more "right" (expectation = 1). The authors report that the 20 *Ss* who verbalized the principle (the number of letters in the word minus one) averaged 42% correct in the block of 9 words *preceding* the block after which correct verbalization occurred, compared with an expectation of 11%. The 28 *Ss* who did not verbalize correctly had a slight but significant rise in their average performance curve. However, Philbrick and Postman point out that there were hypotheses which were partly related to word length, and in their report they do not eliminate the possibility that the above-chance performances can be accounted for by the *Ss*' awareness of these related hypotheses.

D. *Experiments in Which the Subject Does Not Believe That a Cue Is Being Given*

1. *Behavior Not Learned During the Experiment*

In 1900 Dunlap (**11**) reported that the Muller-Lyer illusion had been obtained with 4 *Ss* by making the obliques merely subliminal shadows, so that *S* thought he was simply bisecting a horizontal line segment. Titchner and Pyle (**69**) replicated this experiment with negative results. Manro and Washburn (**43**) also reported negative results, though their paper includes some positive evidence. Both Hollingworth (**28**) and Bressler (**3**) reported positive results; since 1931 the controversy seems to have been dropped.

In 1939 Miller (**47**) reported an experiment in which *S* guessed (by "telepathy") geometric figures while staring into a mirror ("as into a crystal ball") on the back of which the figures were being projected at very low intensities without *S*'s knowledge (these were the "naive" *Ss* in the experiment referred to in Section I, A). At the highest intensity the 5 *Ss* averaged 45% correct, compared with expectation of only 20%; one *S* achieved 91% correct. These *Ss* all showed surprise when told that the figures were actually being projected. No replication of this part of Miller's experiment has been reported.

Perky's classic study (**53**) in which *Ss* were asked to draw various objects which they were to imagine on a ground-glass screen, but which were actually being projected, gave clear-cut evidence of discrimination without awareness, but has apparently never been replicated, except insofar as the studies by Miller can be considered a replication. The stimuli

used in Perky's experiment, however, were clearly supraliminal for *S*s who knew that images were being projected; various precautions were taken to prevent naive *S*s from becoming aware, such as making the edges of the images uneven, moving the image to and fro slowly, and distracting the subject and removing the image as soon as the *S* began his description.

Sidis (**60**) gave each of 20 *S*s the task of reproducing a complicated drawing, then presented a piece of cardboard from which *S* was to choose one of the digits 26471538 "to break up attention." On the margin was written one of the digits. The *S*s did not notice that the digits were on the margin, but tended to choose them 39.4% of the time (expectation = 12.5%).

Coover (**9**) exposed tachistoscopically a card with a capital in the lower right corner, a digit in the upper left corner. *S* was told only to look for a capital letter in the lower right corner, and then to put down any digit that should "come to mind." With 26 *S*s the percentage choices coinciding with the upper left digit varied from 9.5% to 36.7%, with a total of 15.4% (expectation = 10%). Coover reported that 15 of the *S*s showed the effect. However, in view of Coover's statement that the *S* *usually* did not know a digit was being presented with the letter, these results are rather inconclusive as to discrimination without knowing that a cue is being given.

Collier (**8**) repeated Coover's experiment, with some modifications. A figure similar to a Landolt Circle was exposed tachistoscopically at the fixation point and one of six geometric forms was exposed simultaneously in the periphery. *S* had to guess the position of the break in the circle and to name one of the six geometric forms—"the first one that comes to mind." Even though they did not know a form was being presented, the 10 *S*s gave the peripheral figure on 22.2% of the trials, compared with expectation of 16.7%. This was the largest of Collier's results, with peripheral figures of 10 minutes of arc; figures of 15 minutes and 30 minutes gave smaller results.[10]

2. *Response Learned or Strengthened During the Experiment*

In a second study by Miller (**48**), using the same apparatus and general procedure as the one previously described, rewards and punishments were varied among four groups and differences in improvement were obtained. The group showing the greatest improvement gave 63% correct compared with an expectation of 50% (only two figures were used with this group); all 4 *S*s in the group are reported to have learned.

In his experiments on extrasensory perception and related phenomena Kennedy (**34**) found that one *S* among 100 college students tested with the "open matching method" made 275 hits in 1,000 matches (expectation = 200; $p < .00001$). *S* was allowed to look closely at the backs of the cards and to tilt them to get reflections; he reported throughout that he was not aware of using visual cues; however, when questioned about his close scrutiny and tilting of the cards he replied that this was necessary or "there would be nothing to go on" (**34**, p. 150). Kennedy con-

[10] Collier relates this finding to a result obtained by Coover that discrimination of digits 8 degrees in the periphery was better than at 6 degrees or 4 degrees. The percentages right in Coover's experiment were 15%, 12%, and 13% respectively. At 2 degrees the percentage went up to 22%, but the *S*s became aware.

cludes, "The important point is that they (the cues) may be used subliminally by naive *S*s who glance at the backs of the cards" (**34**, p. 150). Few psychologists would disagree with this conclusion, but the fact remains that none of Kennedy's 100 *S*s gave clear-cut instances of this phenomenon; the phenomenon used in the explanation is almost as difficult to obtain in the laboratory as ESP itself.

In 1932 Thorndike reported in the *Fundamentals of Learning* (**67**) a number of experiments[11] dealing with learning without awareness, under the chapter title, "After-effects without revival." Thorndike's major purpose in these experiments was to eliminate the possibility of rehearsal (or "inner repetition") and thus to have the opportunity of showing that the aftereffects of satisfyingness can strengthen a connection even without the advantage of frequency. In his discussion of the results, however, he emphasized also the direct "confirming" influence of the reward as opposed to an effect mediated by ideas (awareness). This position of Thorndike is well known; it should be distinguished from the assertion that a given kind of behavior can be *maintained at a high consistency* without the individual's becoming aware. He would presumably have made the latter assertion with respect to many kinds of behavior (e.g., playing tennis, recognizing faces), but to this reviewer's knowledge he did not discuss the range of behaviors with respect to which it would apply. He was aware that the effects of learning without awareness which he obtained in the laboratory were, in general, slight.

Thorndike reported several experiments involving stimulus discrimination. In one of these, *S* completed mutilated words, each of which could be completed in more than one way. During training *E* said "Right" when a dot following an *a* was filled by a *v*, a dot following a *b* was filled by an *l*, etc., and "Wrong," otherwise. Thorndike reported that 3 of 8 *S*s learned to respond without awareness of the rules; the data which he presents support the conclusion for the training series (which can be accounted for by memory of specific words, as the same words were presented at least 14 times) but not for the initial vs. final test trials.[12] Another procedure used by Thorndike was to present *S* with a list of rare words (in one experiment, Spanish words) and instruct him to underline a synonym for each, from five possible choices. *E* announced "Right" or "Wrong" after each choice. The words were so arranged that in each 100 trials the correct word occurred most frequently in the first place to the right, less frequently in the second place, . . . , least frequently in the first place to the left. The *S*s were questioned at the end, and only those unaware that position had any relation to correctness were included in the analysis of results. The result upon which Thorndike relied most was a test of position choices before vs. after training. Choices of positions 1 and 2 (from the left) declined from 43% to 37%; choices of positions 4 and 5 increased from 38% to 43%. In an experiment with Spanish words, comparable figures are a decline from 52% to 40%

[11] Some of these did not involve stimulus discrimination and will be discussed in Section II.

[12] Thorndike states that the initial vs. final test yielded a significant difference, but the data which he presents do not support this conclusion. Ironically, he obtained a significant initial vs. final difference for a control group (experiment 57) with forced repetitions and no announcement of right or wrong, but nothing is said about the *S*s' awareness.

and an increase from 27% to 41%. An alternative explanation for these results is that the accuracy of the *S*'s discrimination of the rare (or Spanish) words improved with training. This explanation is not only plausible; it is supported to some extent by the data themselves.[13] In other experiments *S* was asked to judge, on each of several hundred trials, which of four lines was longest. Two hundred cards were used, of which 20 had numerals in the middle of the card, 20 had numerals at the right hand end, 20 had one of the lines drawn two millimeters wider than the other, etc. The actual differences in length were "imperceptible," but *E* put the identifiable features of the cards into 100% relation with the position of the line, announcing which line was "right" after each trial. *S* began by trying to see the difference, but usually gave up careful scrutiny after the first few trials and began to guess quickly. In some experiments, however, the differences in length were made somewhat greater, so that *S* could actually get about 44% right without regard to identifiable features of the cards. Thorndike reports increases in the percentage of "Right" responses, without awareness of the use of the cues; however, the results are reported rather sketchily and are difficult to evaluate, as most of these experiments do not meet present methodological standards, particularly with respect to controversial topics. There are three main objections:

1. The possibility is not always eliminated that an improvement in discrimination of "imperceptible" differences has not occurred (which would at best place some of these experiments into Section I, A, 2).

2. Evidence of lack of awareness is often not sufficiently complete.

3. The possibility is not always eliminated that *S*s simply learned to respond to particular cards.

Experiment 61 (pp. 239–240), with cards and procedure as described above, seems to be most nearly free of any valid criticism.[14] After training, *S*s averaged 33.1% correct choices compared with 25.1% before training (expectation = 25%).

Thorndike and Rock (68) told *S* to respond to a word given by *E* with the first word that he thought of, and that *E* would say "Right" when the response word was any of 10 or more arbitrarily selected words, and that he would receive a money bonus based both on speed and on the number "Right." Actually *E* said "Right" if the response word was "clearly due to sequential connections used in speaking or writing" and "Wrong" if the response was "clearly due to connections used in getting the word's meaning." As the *S*s were not questioned at the end of the experiment, it is impossible to say how many of them assumed that some cue was being given. Thorndike and Rock reported that, of the 23 *S*s who showed some learning (of a total of 30), only

[13] This explanation helps to account for the different results for *S*s AIII and BI and the failure of BI and BII to show the tendency to underline toward the right in experiment 50. The data on page 214 show considerable discrimination during training, and Thorndike indicates in a footnote on page 225, in connection with another procedure, that the *S*s did have some knowledge of the meanings of rare English words. The results of the "subtle" analysis of experiment 49 (pp. 215–216) can be accounted for by the tendency to drop out wrong responses *or* by increased discrimination.

[14] The only question is whether "Right" was called as often for line 1 as for line 2, etc.; if not, the performance of the *S*s may have been a case of "probability discrimination" similar to that studied in Thorndike's experiments discussed in Section II, B.

one *S* may have become aware. The criteria used for awareness, however, were suddenness of discrimination and 100% responding. Irwin, Kaufman, Prior, and Weaver (**32**) showed that these criteria are not adequate, as their *S*s continued to improve after being taught the distinction and some did not attain a high level of discrimination.

The replication of this experiment by Postman and Jarrett has been discussed in Section I, C, as it is clear in the Postman and Jarrett procedure that *S* knew that some cue was being given.

Cohen, Kalish, Thurston, and Cohen (**6**) used as *S*s 40 male ambulatory patients, all under 55, from a general medical population, excluding neurology and psychiatry. Each *S* was shown 80 cards, each containing one verb and six pronouns, and was told to make up a sentence beginning with one of the pronouns and including the verb. *E* said "Good" after each sentence which *S* began with "I" or "We," beginning with the second block of 20 cards. The authors report that questioning of *S*s at the end of the series revealed no awareness of the contingency between their responses and those of *E*, although their use of "I" and "We" increased from 44% on the first 20 trials to 64% on the last 20, compared with a control group's change from 43% to 41%. In a second experiment the increase was from 41% to 55%.[15]

[15] The authors state that each *S* was asked which pronoun he had used most, and then asked whether *E*'s behavior influenced his choice of a pronoun. It is reported that the *S*s of experiment I were not aware of the contingency between their responses and those of *E*, but no information is given about the *S*s of experiment II in this respect, nor is it reported whether the *S*s in either experiment were aware that they used "I" and "We" more often than the other pronouns.

Klein (**36**), using essentially the same procedure as Cohen et al., obtained both conditioning and extinction without awareness. His *S*s were 80 neuropsychiatric patients in one experiment and 30 in another. Taffel (**66**) used this procedure with 90 psychotic and neurotic male patients and obtained differences in conditioning which were directly related to scores on the Taylor manifest anxiety scale. The "high" anxiety group increased their "I" and "We" responses from 41% on the first 20 trials to 76% on the fourth 20 trials, the "medium" group from 42% to 69% and the "low" group from 39% to 46% (not significant).

In none of the last three papers is there much information given about what the *S*s thought they were doing, nor is the problem of awareness of correlated hypotheses discussed.

Positive results of conditioning to stimuli presented to hysterically anesthetic areas and not consciously perceived by *S* have been reported by Sears and Cohen (**59**) and by Cohen, Hilgard, and Wendt (**7**). Sears and Cohen established a CR to a quick brush of a whisp of cotton across the back of the hysterically anesthetic left hand of a patient, though the patient denied feeling anything. Shock was used as the S_{unc} and the CR was obtained by first establishing a CR to a sharp rap with a pencil, which the patient could feel. The CR to the cotton extinguished in one trial, then, after reconditioning, in two trials. After questioning of the patient and denial on her part that she had felt the cotton, further trials evoked both the CR and conscious awareness of the stimulus. The anesthesia disappeared completely at this time and had not returned six months later. Sears and Cohen had been unable to obtain a CR by simply pairing the

cotton with shock, or by first establishing a CR of the right (nonanesthetic) hand.

In the study by Cohen, Hilgard, and Wendt it was shown first that eyelid reflexes to sound could be very consistently reinforced or inhibited by presenting a light to the blind area either .045 sec. (for reinforcement) or .225 sec. (for inhibition) before the sound was presented. Not once did the *S* report the light. A conditioned response to light in the blind area was then established by using a puff of air as the S_{une}. Next a conditioned verbal response to light was established[16] by pairing light with a sound to which the patient had been trained to shout "Light." This last CR was slow to develop and very weak, perhaps understandably, as the patient saw nothing to shout about.

In his studies of salivary conditioning Razran (**57**) has shown that both conditioning and generalization can occur with no awareness that more salivation is elicited by the S_e and by closely related (phonetographically or semantically) stimuli than by other stimuli. Making the *S*s aware of what was happening or having them adopt conscious facilitory or inhibitory attitudes seemed to have little effect.

Diven (**10**) instructed his *S*s to respond to each stimulus word with the first word that came to mind, continuing to call words until told to stop. The *S*s were also told, "Don't try to make sense of the task," (**10**, p. 295). Forty stimulus words were given, with 12 sec. between each stimulus word and "stop." A 1-sec. shock was administered after the word "barn," which was presented six times during the list. After five minutes *S* was asked to recall the list, then the list was presented without shock, then *S* was again asked to recall. Several groups were run, with varying periods between the first and second sessions. Diven reported that 21 of 52 *S*s remained unaware of the fact that "barn" was the signal for shock, although their physiological indices of anxiety in response to this stimulus had increased. He also reported that the increment in the index due to an increased time interval between sessions was greater for the unaware *S*s than for the aware *S*s, and that the unaware *S*s extinguished less readily.

Haggard (**17**), utilizing Diven's technique, instructed his *S*s to respond to each stimulus word with related words. After 10–12 sec., *E* either gave a strong shock or said "stop." Forty-two stimulus words were given, of which 5 were the word "sword," 5 were "sharp," 10 were additional war words, 10 were peace words, 5 were buffer words following "sword," and 7 were neutral words. *S* was shocked only after "sword." The word "sharp" always preceded "sword." At the end of the conditioning session *S*s were asked whether they could predict the shock; 7 (of 18) were unaware when it came, but were certain it had not come after any given word.[17] There was a significant increase in GSR to the precritical word "sharp" by the unaware group; this increase was 2.6 times that shown by the aware group.

Lacey and Smith (**37**), pointing out certain flaws in Diven's and Haggard's experiments, told *S*s to chain

[16] A previous study by Hilgard and Wendt (**27**) had failed to obtain this effect in a case of hemianopsia following surgical resection of the left occipital lobe.

[17] Haggard (**17**) presumably means *stimulus* word but does not say so explicitly. Two *S*s, who were uncertain whether shock came after "sword" or some other war word, were placed in the unaware group "for convenience in treating the data" (**17**, p. 261).

associate to each stimulus word until *E* said "stop" (after 15 sec.) and to tap meanwhile on a telegraph key at an even rate. In Group I a strong 5-sec. electric shock was given within .5 sec. after "stop" when *S* had been chain associating to "cow." In Group II the shock was given after *S* had been chain associating to "paper." "Cow" and "paper" were each presented 6 times; of the remaining 28 words, 14 were rural and 14 were nonrural words (Diven had reported generalization to rural words). The measure of conditioning was increase in heart rate during the 15 sec. of associating over the previous 15 sec. (maximum minus maximum). Of 31 *S*s, 22 were unaware of the fact that the shock followed chain association to a certain word; of these, 10 were in Group I and 12 were in Group II. Seven of the 10 unaware *S*s in Group I and 8 of the 12 unaware *S*s in group II showed a greater heart rate during association to the critical word than association to noncritical words. Group I unaware *S*s also showed generalization to other rural words in the list. In a further analysis Lacey, Smith, and Green (38) showed that the aware *S*s immediately developed a strong emergency reaction that did not increase with the number of reinforcements but instead showed a gradual adaptation. Unaware *S*s showed typical conditioning curves, at a much lower level of autonomic activity and discrimination than the aware *S*s. It was also shown that the unaware *S*s had generalized more than the aware *S*s. These very interesting and important results have not, to the reviewer's knowledge, been replicated.

Eriksen and Kuethe (13) instructed their 31 *S*s to give an association to each stimulus word as quickly as possible, and that they would be shocked if their association was too slow for that particular stage of practice and also under a second condition which they might discover during the session and thus avoid the shock. Five of the 15 stimulus words were randomly designated as critical; on the first trial the *S*s' associations to these five words were followed by a painful electric shock to the ankle. On succeeding trials *S* was shocked whenever he repeated one of the originally shocked associations, until a total of 10 trials or two successive shockless trials had occurred. The *S* was assured that there would be no more shock and was then asked to give chained associations to the stimulus words, being allowed 15 seconds for each chain. Finally *S* was interviewed to determine the level of awareness of the method of avoiding shock. Eleven *S*s could be classified as having a high level of awareness and another 11 as having a low level of awareness; the latter were unable to state any reasons for the shock other than the misinformation given them at the beginning. The low-level group, however, showed a drop in repetitions from 42% on the second trial to only 13% on the seventh trial for the critical words compared with 63% and 60% respectively for the noncritical words; their avoidance behavior in terms of percentages was not significantly different from the high-level group. An additional finding of considerable interest in this experiment was that although the low-level group showed even greater avoidance on the second and third trials than the high-level group, their reaction times tended to drop, whereas the reaction times of the high-level group tended to rise. Conscious avoidance seemed to take longer than automatic, low-level avoidance. Both groups gave about

50% of the first-trial responses to critical stimuli during the chained associations; this compared with about 70% of the first-trial responses to noncritical stimuli.

Hildum and Brown (**25**) conducted an attitude survey over the telephone, asking each *S* 15 questions concerning the Harvard philosophy of General Education (sic). One group received "Good" after "pro" answers, another "Good" after "anti" answers, another "Mm-hmm" after "pro," and a fourth "Mm-hmm" after "anti." The "Good—pro" group gave significantly more favorable responses than the "Good—anti" group, whereas the difference for the "Mm-hmm" groups was not even in the predicted direction. Eight of the 20 "Good" *S*s were aware that *E* had said "Good" whereas only one of the 20 "Mm-hmm" *S*s was aware of "Mm-hmm." The authors, who show considerable awareness of the problem of accurate determination of awareness, point out that awareness of *E*'s behavior may in some way account for the greater effectiveness of "Good," even though the denial, by all the *S*s, that they were influenced by *E* be accepted at face value. The question might be raised whether the *S*s, who were Harvard undergraduate and graduate students, may have felt that they should not have been influenced; if so, the denial should probably not be taken at face value.

Nuthmann (**51**) presented *S*s with a personality test containing 100 acceptance-of-self items and 40 buffer items. Each item was presented on a card; *S* then pressed a "true" or a "false" switch. The 45 *S*s had all had the test previously and had scored in the lowest 15% of a group of 420 students. In one group an acceptance-of-self response (which could be either "true" or "false," depending upon the item) was followed by "Good"; in another group a light came on; in the control group no reinforcement was given. The 15 "Good" *S*s increased from about 67% acceptance-of-self responses during the first 20 critical trials to about 77% such responses during the last 20 critical trials (each block contained an additional 8 buffer-item trials), significantly higher than the light and control groups, which declined slightly from about 65% to about 64% acceptance-of-self responses. The "Good" *S*s were asked, "Did you think that anything I said had anything to do with your responses?" Five verbalized the purpose of the experiment correctly; these had responded at a slightly but insignificantly higher level throughout than the unaware *S*s; four of the five volunteered the information that they were quite sure that they had not been influenced by the reinforcement. It is unfortunate that a comparison is not presented between the unaware "Good" *S*s and the control *S*s; the author points out that not all the *S*s in the "Good" group increased in acceptance-of-self responses over the five blocks. Correlated hypotheses are mentioned to the extent of stating that most of the unaware *S*s verbalized a "contradictory" hypothesis, many stating that reinforcement occurred whenever their response was consistent with that of the previous testing. Differential memory effects may have entered in to make this hypothesis less contradictory than it seemed. Four of the "Light" *S*s verbalized the purpose correctly, yet did not show a significant increase in acceptance-of-self responses.

In the Rees and Israel experiment (**58**) referred to in Section I, B, *S*s were given a training series of 15 anagrams and then a critical series of 15.

In the critical series each anagram had two solutions, one being the rearrangement of letters according to the scheme 54123. The training series for the 10 experimental *S*s consisted of anagrams which had only the solution 54123; the training series for the 10 control *S*s had various solutions. Six of the 10 experimental *S*s had no idea that the solutions involved any regular order of rearrangement; yet about 90% of their solutions were 54123, the comparable figure for the control group being about 50%. This experiment is usually discussed under the heading of "set"; what seems to be involved is the establishment of visual-motor behavior without conscious awareness by *S*, perhaps similar to the complex visual-motor behaviors involved in reading and in many sports.

II. Behavior Not Involving Stimulus Discrimination

In experiments not involving stimulus discrimination, in place of the variable of whether *S* believes a cue is being given we have the variable of whether *S* believes that his task is to discover a method of responding. With one exception all the experiments reviewed in this section depend upon learning during the experiment.

A. *Experiments in Which* S *Believes That His Task Is to Discover a Method of Responding*

Sidowski (**61**), using a procedure similar to that of Greenspoon (**16**), instructed *S*s to say any words they could think of, then pointed out a light bulb mounted in front of *S* which could be made to blink, telling *S* it was his job to make the bulb blink as many times as possible during the experiment. Each plural word given by *S* was followed by a blink. After 325 responses each *S* was asked, "Were you aware of the purpose of the experiment?" and "Were you aware of the purpose of the light?" Only 3 of 13 *S*s stated that the purpose was to condition plural words, and the remaining 10 showed an increase from about 23% plurals during the first 65 words to about 35% during the last 65 words. However, Sidowski does not discuss the problem of correlated hypotheses.

Weiss (**75**) instructed his *S*s to respond to each word from *E* with a word or phrase; *S* was told that *E* would say "Right" or "Wrong" according to whether *S*'s word fell into a certain category. The category chosen was that of living things. Neither a timed group ("respond as quickly as you can") nor an untimed group differed significantly from a control group (untimed with no reinforcement). The use of discrete trials may be an important difference from Sidowski's and Greenspoon's experiments.

B. *Experiments in Which* S *Does Not Believe That His Task Is to Discover a Method of Responding*

1. *Response Not Learned During the Experiment*

Kennedy (**33**), replicating an experiment by Hansen and Lehmann (**19**), got *S*s to whisper without awareness that they were doing so. Each *S* was given kinesthetic ("mental shouting") instructions or "visualization" instructions; in both cases unconscious whispering occurred.

2. *Response Learned or Strengthened During the Experiment*

In the same chapter of *Fundamentals of Learning* referred to in Section I, D, Thorndike reported several experiments not involving stimulus discrimination. His basic procedure was

to present *S* with a series of slips of paper, asking him to estimate the length of each to the nearest quarter-inch. Certain lengths were presented more frequently than others. *E* announced "Right" or "Wrong" after each guess. A similar procedure was used for the judgment of areas. At the conclusion of his judgments, each *S* was asked whether he had thought that any lengths occurred more often than others, and, if so, which ones. Thorndike found that those lengths or areas which were right more frequently began to be given more frequently even to slips of paper of some other magnitude. These results, however, can be accounted for either by the learning of specific cards (in some experiments) or by improved discrimination of magnitudes.

Greenspoon instructed *S*s to say all the words they could think of, to say them individually, and not to use any sentences or phrases (**16**). By saying "mmm-hmm" after each plural response, during the first 25 minutes, he more than doubled the percentage of plurals for 14 *S*s, compared with a control group which had no contingent response from *E*. When questioned at the end, none of the 14 *S*s verbalized the relation between the contingent response and the response which it followed (one *S* had been dropped because he did verbalize the relation), although presumably the *S*s were aware that *E* had been saying "mmm-hmm." With another group, the contingent response followed nonplural responses, but did not increase them. Similarly, "huh-uh" decreased plurals but not nonplurals. Greenspoon points out that the differential effect may depend upon size and/or heterogeneity of the response class. Greenspoon does not discuss the problem of correlated hypotheses.

Sidowski (**61**), in the same experiment referred to in Section II, A, had one group of *S*s who were given the light-blink without any mention of it in the instructions. Only 3 of the 19 *S*s in this group verbalized the purpose of the experiment correctly, yet the remaining 16 showed an increase in plurals from about 15% during the first 65 words to a maximum of about 27% during the third block of 65 words, significantly better than control groups without reinforcement of plurals. As with the group previously discussed, no mention is made of the problem of correlated hypotheses.

Verplanck has reported an operant conditioning experiment in which various *E*s increased the rate of opinion-stating of their *S*s by reinforcing in various ways (repetition, agreement, etc.) (**71**). Verplanck reports that no *S* gave any evidence that he was "aware" that he was serving as a subject in an experiment, that his behavior was being deliberately manipulated and recorded, or that he recognized that there was anything peculiar about the conversation. To come within the scope of behavior without awareness, however, as dealt with in this review, it would have to be shown that the *S*s were not aware that their rate of stating opinions increased and/or that their opinion-stating was being influenced by *E*'s responses (this lack of awareness is not necessarily covered by the statement that they were unaware that their behavior was being *deliberately* manipulated). Elsewhere, Verplanck has indicated that behavior without awareness has been obtained, but has not yet published his results in detail (**72**).[18]

[18] The reviewer, while serving as *S* for Verplanck in an informal discrimination experiment with "trading" cards, apparently began to use the color blue as a cue without being aware of using it; this result, which was quite convincing and startling to the reviewer, is unfortunately not easily replicated.

TABLE 2

STUDIES INCLUDED IN THIS REVIEW

Replications of preceding experiments indicated by (R); positive results by (+); negative results by (−); alternative explanations by (?).

I. Behavior Involving Stimulus Discrimination	Behavior Not Learned During Experiment	Behavior Learned or Strengthened During Experiment
A. *S* knows specific nature of cue	Peirce & Jastrow (+) Fullerton & Cattell (+) Sidis (+) (R) Stroh, Shaw, & Washburn (+) (R) Pillai (+) Coover (+) Baker (+) (?) (R) Hilgard, Miller, & Ohlsen (−) Williams (+) (R) Miller (+) (R) Vinacke (−)	Newhall & Sears (+) (?) Baker (+) (?) (R) Wedell, Taylor, & Skolnick (−) (R) Hilgard, Miller, & Ohlsen (−) Lazarus & McCleary (+) (?)
B. *S* knows general nature of cue		Hull (+) (?) Smoke (+) (?) Heidbreder (+) (?) Goodnow & Postman (+) (?)
C. *S* believes cue is being given		Postman & Jarrett (+) (?) Philbrick & Postman (+) (?)
D. *S* does not believe cue is being given	Dunlap (+) (?) (R) Titchner & Pyle (−) (R) Manro & Washburn (±) (R) Hollingworth (+) (R) Bressler (+) Perky (+) Miller (+) Sidis (+) Coover (?) (R) Collier (+)	Miller (+) Kennedy (+) (?) Thorndike (+) (?) Thorndike & Rock (+) (?) Cohen, Kalish, Thurston, & Cohen (+) (?) Klein (+) (?) Taffel (+) (?) Sears & Cohen (+) Cohen, Hilgard, & Wendt (+) Razran (+) Diven (+) (?) Haggard (+) (?) Lacey & Smith (+) Eriksen & Kuethe (+) Hildum & Brown (+) (?) Nuthmann (+) (?)
II. Behavior Not Involving Stimulus Discrimination		
A. *S* believes task is to discover method of responding		Sidowski (+) (?) Weiss (−)
B. *S* does not believe task is to discover method of responding	Hansen & Lehmann (+) (R) Kennedy (+)	Thorndike (+) (?) Greenspoon (+) (?) Verplanck (+) (?) Mandler & Kaplan (+) (?) Rees & Israel (+).

Mandler and Kaplan (**42**) ran 28 *S*s with the Greenspoon procedure, saying "mm-hmm" after every plural-noun response during the reinforcement period. Each *S* gave 500 responses, the first 100 being used to obtain the operant level, the last 200 being extinction trials. An extended interview given to each *S* was used to determine the degree of awareness of the reinforcing contingency, on a rating scale of from 1 (full awareness) to 11 (complete lack of awareness). The ratings ranged from 3 to 11 with a mean of 6. The *S*s as a group did not show a significant increase in plural noun responses, but those *S*s who thought (as determined by the interview) that "mm-hmm" meant they were doing all right increased from about 12% operant level to about 19% during the first reinforcement period (100 trials) whereas those who thought that "mm-hmm" meant they were not doing all right declined from about 12% to 7%, even though their mean awareness was 5.6 compared with a 6.4 for the "positive" group! This difference between the two groups in plural-noun responses is significant at the .01 level. During the second reinforcement period, however, the "positive" group fell down to about 9%, not significantly different from the "negative" group. The fact that many *S*s were to some extent aware plus the problem of correlated hypotheses makes these results difficult to evaluate.

Summary and Conclusions

Table 2 presents a classification of the studies on behavior without awareness covered by this review. Positive results are indicated by "(+)," negative results by "(−)," replications of preceding experiments by "(R)," and studies for which alternative explanations have been suggested, either by this reviewer or by others (in some cases the authors themselves) or which do not present conclusive evidence are indicated by "(?)." If we remember that the results in I, D tend to be small and/or unreplicated, then it is clear from the table that, contrary to a widespread conviction among psychologists, the only type of behavior without awareness which can be easily reproduced on the basis of published reports is the classical type, in which *S* knows *what* he is supposed to be discriminating, but does not know *that* he is discriminating, because of the absence of the usual sensory experiences to which he is accustomed under the given type of stimulation. Nevertheless we know (in some sense of the word) that many other kinds of behavior without awareness do frequently occur, and psychologists continue to report their demonstration in the laboratory. The establishment of conditions under which these other types can be unequivocally demonstrated by and measured by any competent researcher remains an interesting and important challenge to experimental ingenuity.

References

1. Baker, L. E. The influence of subliminal stimuli upon verbal behavior. *J. exp. Psychol.*, 1937, **20**, 84–100.
2. Baker, L. E. The pupillary response conditioned to subliminal auditory stimuli. *Psychol. Monogr.*, 1938, **50**, No. 223.
3. Bressler, J. Illusion in the case of subliminal visual stimulation. *J. gen. Psychol.*, 1931, **5**, 244–250.
4. Bricker, P. D., & Chapanis, A. Do incorrectly perceived tachistoscopic stimuli convey some information? *Psychol. Rev.*, 1953, **60**, 181–188.
5. Cason, Hulsey, & Katcher, Naomi. An attempt to condition breathing and

eyelid responses to a subliminal electric stimulus. *J. exp. Psychol.*, 1934, **16**, 831–842.

6. COHEN, B. D., KALISH, H. I., THURSTON, J. R., & COHEN, E. Experimental manipulation of verbal behavior. *J. exp. Psychol.*, 1954, **47**, 106–110.
7. COHEN, L. H., HILGARD, E. R., & WENDT, G. R. Sensitivity to light in a case of hysterical blindness studied by reinforcement-inhibition and conditioning methods. *Yale J. Biol. Med.*, 1933, **6**, 61–67.
8. COLLIER, R. M. An experimental study of the effects of subliminal stimuli. *Psychol. Monogr.*, 1940, **52**, No. 5 (Whole No. 236).
9. COOVER, J. E. Experiments in psychical research. *Psychical Research Monograph No. 1.* Stanford: Stanford Univer. Press, 1917.
10. DIVEN, K. Certain determinants in the conditioning of anxiety reactions. *J. Psychol.*, 1937, **3**, 291–308.
11. DUNLAP, K. Effect of imperceptible shadows on the judgments of distance. *Psychol. Rev.*, 1900, **7**, 435–453.
12. ERIKSEN, C. W. Subception: fact or artifact? *Psychol. Rev.*, 1956, **63**, 74–80.
13. ERIKSEN, C. W., & KUETHE, J. L. Avoidance conditioning of verbal behavior without awareness: a paradigm of repression. *J. abnorm. soc. Psychol.*, 1956, **53**, 203–209.
14. FULLERTON, G. S., & CATTELL, J. MCK. *On the perception of small differences.* Univer. of Penna. Publ., Philos. Series, 1892, No. 2.
15. GOODNOW, JACQUELINE J., & POSTMAN, L. Probability learning in a problem-solving situation. *J. exp. Psychol.*, 1955, **49**, 16–22.
16. GREENSPOON, J. The reinforcing effect of two spoken sounds on the frequency of two responses. *Amer. J. Psychol.*, 1955, **68**, 409–416.
17. HAGGARD, E. A. Experimental studies in affective processes: I. Some effects of cognitive structure and active participation on certain autonomic reactions during and following experimentally induced stress. *J. exp. Psychol.*, 1943, **33**, 257–284.
18. HAKE, H. W., & HYMAN, R. Perception of the statistical structure of a random series of binary symbols. *J. exp. Psychol.*, 1953, **45**, 64–74.
19. HANSEN, F. C. C., & LEHMANN, A. Ueber unwillkurliches Flustern. *Philosophische Studien*, 1895, 11, 471–530.
20. HEIDBREDER, E. An experimental study of thinking. *Arch. Psychol.*, 1924, **11**, No. 73.
21. HEIDBREDER, E. The attainment of concepts: III. The process. *J. Psychol.*, 1947, **24**, 93–138.
22. HEIDBREDER, E. The attainment of concepts. VIII. The conceptualization of verbally indicated instances. *J. Psychol.*, 1949, **27**, 263–309.
23. HEIDBREDER, E., BENSLEY, M. L., & IVY, M. The attainment of concepts: IV. Regularities and levels. *J. Psychol.*, 1948, **25**, 299–329.
24. HEIDBREDER, E., & OVERSTREET, P. The attainment of concepts: V. Critical features and contexts. *J. Psychol.*, 1948, **26**, 45–69.
25. HILDUM, D. C., & BROWN, R. W. Verbal reinforcement and interviewer bias. *J. abnorm. soc. Psychol.*, 1956, **53**, 108–111.
26. HILGARD, E. R., MILLER, J., & OHLSON, J. A. Three attempts to secure pupillary conditioning to auditory stimuli near the absolute threshold. *J. exp. Psychol.*, 1941, **29**, 89–103.
27. HILGARD, E. R., & WENDT, G. R. The problem of reflex sensitivity to light studied in a case of hemianopsia. *Yale J. Biol. Med.*, 1933, **5**, 373–385.
28. HOLLINGWORTH, H. L. *Advertising and selling.* New York: Appleton, 1913.
29. HOWES, D. A statistical theory of the phenomenon of subception. *Psychol. Rev.*, 1954, **61**, 98–110.
30. HOWES, D., & SOLOMON, R. L. A note on McGinnies' "Emotionality and perceptual defense." *Psychol. Rev.*, 1950. **57**, 235–240.
31. HULL, C. L. Quantitative aspects of the evolution of concepts: An experimental study. *Psychol. Monogr.*, 1920, **28**, No. 1 (Whole No. 123).
32. IRWIN, F. W., KAUFMAN, K., PRIOR, G., & WEAVER, H. B. On "Learning without awareness of what is being learned." *J. exp. Psychol.*, 1934, **17**, 823–827.
33. KENNEDY, J. L. Experiments on "unconscious whispering." *Psychol. Bull.*, 1938, **35**, 526(a).
34. KENNEDY, J. L. The visual cues from the backs of the ESP cards. *J. Psychol.*, 1938, **6**, 149–153.
35. KENNEDY, J. L. A methodological review of extrasensory perception. *Psychol. Bull.*, 1939, **36**, 59–103.
36. KLEIN, S. Conditioning and extinction of operant verbal behavior in neuropsychiatric hospital patients. *Dissert.*

Abstr., 1954, **14**, 2127–2128. (Abstract of Ph.D. thesis, 1954, Indiana Univer.)
37. LACEY, J. I., & SMITH, R. L. Conditioning and generalization of unconscious anxiety. *Science*, 1954, **120**, 1045–1052.
38. LACEY, J. I., SMITH, R. L., & GREEN, A. Use of conditioned autonomic responses in the study of anxiety. *Psychosom. Med.*, 1955, **17**, 208–217.
39. LAZARUS, R. S. Subception: fact or artifact? A reply to Eriksen. *Psychol. Rev.*, 1956, **63**, 343–347.
40. LAZARUS, R. S., & MCCLEARY, R. A. Autonomic discrimination without awareness: a study of subception. *Psychol. Rev.*, 1951, **58**, 113–122.
41. LEEPER, R. Cognitive processes. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 730–757.
42. MANDLER, G., & KAPLAN, W. K. Subjective evaluation and reinforcing effect of a verbal stimulus. *Science*, 1956, **124**, 582–583.
43. MANRO, H. M., & WASHBURN, M. F. Effect of imperceptible lines on judgment of distance. *Amer. J. Psychol.*, 1908, **19**, 242–243.
44. MCCLEARY, R. A., & LAZARUS, R. S. Autonomic discrimination without awareness: an interim report. *J. Pers.*, 1949, **18**, 171–179.
45. MCGINNIES, E. Emotionality and perceptual defense. *Psychol. Rev.*, 1949, **56**, 244–251.
46. MCGINNIES, E. Discussion of Howes' and Solomon's note on "Emotionality and perceptual defense." *Psychol. Rev.*, 1950, **57**, 229–234.
47. MILLER, J. G. Discrimination without awareness. *Amer. J. Psychol.*, 1939, **52**, 562–578.
48. MILLER, J. G. The role of motivation in learning without awareness. *Amer. J. Psychol.*, 1940, **53**, 229–239.
49. MILLER, J. G. *Unconsciousness*. New York: Wiley, 1942.
50. NEWHALL, S. M., & SEARS, R. R. Conditioning finger retraction to visual stimuli near the absolute threshold. *Comp. Psychol. Monogr.*, 1933, **9**, No. 43.
51. NUTHMANN, ANNE M. Conditioning of a response class on a personality test. *J. abnorm. soc. Psychol.*, 1957, **54**, 19–23.
52. PEIRCE, C. S., & JASTROW, J. On small differences of sensation. *Mem. nat. Acad. Sci.*, 1884, **3**, 73–83.
53. PERKY, C. W. An experimental study of imagination. *Amer. J. Psychol.*, 1910, **21**, 422–452.
54. PHILBRICK, E. B., & POSTMAN, L. A further analysis of 'learning without awareness.' *Amer. J. Psychol.*, 1955, **68**, 417–424.
55. PILLAI, R. P. B. K. A study of the threshold in relation to the investigations on subliminal impressions and allied phenomena. *Brit. J. educ. Psychol.*, 1939, **9**, 97–98(a).
56. POSTMAN, L., & JARRETT, R. F. An experimental analysis of 'learning without awareness.' *Amer. J. Psychol.*, 1952, **65**, 244–255.
57. RAZRAN, G. Stimulus generalization of conditioned responses. *Psychol. Bull.*, 1949, **46**, 337–365.
58. REES, H. J., & ISRAEL, H. E. An investigation of the establishment and operation of mental sets. *Psychol. Monogr.*, 1935, **46**, No. 6 (Whole No. 210).
59. SEARS, R. R., & COHEN, L. H. Hysterical anesthesia, analgesia, and astereognosis. *Arch. Neurol. Psychiat.*, 1933, **29**, 260–271.
60. SIDIS, B. *The psychology of suggestion*. New York: Appleton, 1898.
61. SIDOWSKI, J. B. Influence of awareness of reinforcement on verbal conditioning. *J. exp. Psychol.*, 1954, **48**, 355–360.
62. SILVERMAN, A., & BAKER, L. E. An attempt to condition various responses to subliminal electrical stimulation. *J. exp. Psychol.*, 1935, **18**, 246–254.
63. SMOKE, K. L. An objective study of concept formation. *Psychol. Monogr.*, 1932, **42**, No. 191.
64. SNYGG, D. The relative difficulty of mechanically equivalent tasks. I. Human learning. *J. genet. Psychol.*, 1935, **47**, 299–320.
65. STROH, M., SHAW, A. M., & WASHBURN, M. F. A study in guessing. *Amer. J. Psychol.*, 1908, **19**, 243–245.
66. TAFFEL, C. Anxiety and the conditioning of verbal behavior. *J. abnorm. soc. Psychol.*, 1955, **51**, 496–501.
67. THORNDIKE, E. L. *The fundamentals of learning*. New York: Teachers College, Columbia Univer., 1932.
68. THORNDIKE, E. L., & ROCK, R. T. Learning without awareness of what is being learned or intent to learn it. *J. exp. Psychol.*, 1934, **17**, 1–19.
69. TITCHNER, E. B., & PYLE, W. H. Effect of imperceptible shadows on the judgment of distance. *Proc. Amer. phil. Soc.*, 1907, **46**, 94–109.
70. URBAN, F. M. *The application of statistical methods to the problems of psychophysics*. Philadelphia: Psychological Clinic Press, 1908.

71. VERPLANCK, W. S. The control of the content of conversation: Reinforcement of statements of opinion. *J. abnorm. soc. Psychol.*, 1955, **51**, 668–676.
72. VERPLANCK, W. S. The operant conditioning of human motor behavior. *Psychol. Bull.*, 1956, **53**, 70–83.
73. VINACKE, W. E. The discrimination of color and form at levels of illumination below conscious awareness. *Arch. Psychol.*, 1942, **38**, No. 267.
74. WEDELL, C. H., TAYLOR, F. V., & SKOLNICK, A. An attempt to condition the pupillary response. *J. exp. Psychol.*, 1940, **27**, 517–531.
75. WEISS, R. L. The influence of 'set for speed' on 'learning without awareness.' *Amer. J. Psychol.*, 1955, **68**, 425–431.
76. WILLIAMS, A. C. Perception of subliminal visual stimuli. *J. Psychol.*, 1938, **6**, 187–199.

Received May 9, 1957.

PSYCHOLOGICAL BULLETIN
Vol. 54, No. 5, 1957

ON THE APPLICATION OF GENETIC EXPECTANCIES AS AGE-SPECIFIC BASE RATES IN THE STUDY OF HUMAN BEHAVIOR DISORDERS

JOHN S. PEARSON AND IRENE B. KLEY
Rochester State Hospital, Minnesota

The observation that abnormalities of behavior are relatively frequent among the children of persons who themselves behave abnormally antedates scientific psychology by a number of centuries. For several decades, argument among scientists has waxed and waned as to whether biological inheritance or social inheritance is primarily responsible for this phenomenon. In the nature-versus-nurture controversy, little notice has been paid to the fact that the tendency of particular abnormalities of behavior to "run in families" might be useful to behavioral scientists regardless of whether the ultimate cause lies in genetic endowment or in environmentally determined experience, and regardless of the magnitude of the correlation between social and biological inheritance. This potential usefulness lies in the fact that individuals in a population with a known and relatively high incidence rate for a particular disorder may be submitted to longitudinal investigation of a kind which would not be economical in samples drawn from the general population. For example, assuming it were possible to select a sample of 100 infants in which it was known that 68 cases of schizophrenia would eventually develop, detailed comparison on a variety of measures over a period of years might yield refined concepts regarding the etiology and prediction of schizophrenia in terms of the crucial differences between the 68 persons who ultimately developed the disease and the 32 who did not. Similarly, in a sample of this kind and batting against a known expectancy of 68%, one might evaluate the efficacy of any desired mental hygiene program in preventing schizophrenic breakdown. In such a sample, we believe longitudinal study of the kind proposed would be practicable. By way of contrast, a random sample selected from the general population to assure an eventual total of 68 cases of schizophrenia for purposes of a similar investigation, would require the investigator to start with something like 8,000 infants. This figure is based upon an estimate of .0085 as the incidence of schizophrenia in the general population. The difficulties involved in a detailed longitudinal investigation with this number of subjects are obvious, whether with predictive or preventive goals in view.

THE BASE-RATES PROBLEM

The reader may recognize in the foregoing an instance of the "base-rates" problem. This problem and its implications for psychologists were discussed by Slater and Slater (**22**) in 1944, and elaborated by Meehl and Rosen (**10**) in 1955. In essence, these authors called attention to the fact, often neglected by investigators seeking to predict behavior from psychometric patterns or cutting scores, that the predictive efficiency of an instrument cannot be judged without knowledge of the base rate or frequency of the variable in question in the population where such predictions, if successful, would be of importance. They point out, for exam-

ple, that in seeking to predict suicides within a psychiatric population or disciplinary discharges among Army recruits, the psychologist and his tests are working in the face of odds which are probably well nigh unbeatable. A minute percentage of "false positives" in the traditional test-validation process, when reflected against the total population of psychiatric patients or against all Army inductees, is likely to render the psycologist wrong much more often than if he had stuck with the base rates and predicted no suicides and no disciplinary discharges.

Eventually, like it or no, we will have to come to grips with the high probability that the base-rates problem applies in the prediction of mental disorder from kind and number of traumatic life experiences, just as it applies in the case of psychometric prediction. Generalizations as to the etiology of mental disorder based upon a posteriori inferences from the life histories of clinical cases likewise are valid only if it can be shown that the incidence of mental disorder is higher among individuals who subsequently undergo similar life experiences than among those who undergo different experiences.

As Meehl and Rosen have remarked, behavioral scientists rarely concern themselves with establishing base rates. Studies of various clinical groups are likely to report the frequency of broken homes, over-protective mothers, perception of parental coitus, etc., in the backgrounds of patients without any indication as to the frequency of similar "traumatic" events in the lives of normal persons. The reader is generally left to surmise that such occurrences are as rare among normals as they are characteristic of patient populations.

Studies like that of Oltman, McGarry, and Friedman (12) in relation to a single and relatively discrete behavioral datum (parental deprivation incident to the broken home) in normal, as well as in abnormal, groups, are rarities in psychological literature. Even less frequent are investigations like that of Renaud and Estess (17) which seek to disclose the totality of pathogenic experiences in the childhood of normal, adequately functioning persons through life-history interviewing and by other means.

Oltman et al. found that groups of normals, schizophrenics, manic-depressives, and other psychotics did not differ significantly with respect to the incidence of broken-home backgrounds. Psychopathic personalities and neurotics did tend to come from broken homes with slightly, but significantly, greater frequency than did the psychotic and normal persons. Renaud and Estess were impressed by the fact that the life histories of their 100 normal men revealed "as many traumatic events and pathogenic factors as are ordinarily elicited in interviews with many psychiatric patients."

If one accepts the results of these studies and takes any reasonable estimate of the frequency of disabling psychiatric abnormality in the entire population, he must perforce conclude that in predicting life-adjustment outcomes it is safest to stick with the base rates. Judging from this evidence, the chances are that a child will function adequately as an adult whether or not he is rejected by his parents, mishandled in his toilet training, abused by his siblings, or whatever.

If subsequent research supports the findings that psychic trauma in childhood is so widespread as to render the early life histories of psychiatric patients indistinguishable from those of a vast number of normals, it will be necessary to re-examine the preva-

lent assumption that unfortunate life experiences are both necessary and sufficient causes of "functional" mental disorder. It will likewise be necessary to re-examine the rationale of therapies and prophylactic measures predicated on this assumption.

The question as to why some individuals emerge from the worst possible socio-biological backgrounds to become and remain productive, well-adjusted adults, while other individuals from good backgrounds become severely neurotic or psychotic in the face of moderate or minimal stress can hardly be answered by methodology which ignores the base rates. In following the suggestion of Meehl and Rosen that we should study subpopulations with base rates higher than obtain in the sampling universe for relatively rare behavior abnormalities, there is much to be said for a closer look at the methodology and findings of human geneticists with respect to mental disorder.

Models for Psychogenetic Research

Huntington's chorea, perhaps the first mental disorder to be thoroughly understood from a genetic point of view, supplies a paradigm for a psychogenetic approach to the base-rates problem, and illustrates a possible means for beating the base rates in longitudinal research on the psychometric detection of organic brain damage.

Psychologists have long been intrigued by the notion that tests of intellectual functioning might be devised which would be of practical importance in distinguishing people with demonstrable brain lesions from other people with "functional" impairment and from normals. To date, efforts to develop such measures have been largely disappointing, due possibly to inherent sampling problems (**24**). If one utilizes psychiatric patients submitted to psychosurgery and compares test performances before and after leucotomy or lobotomy, it is difficult to justify any far-reaching generalizations regarding frontal lobe functions in people who are not sufficiently psychotic to require such surgery. Cases of accidental brain trauma rarely accommodate an investigator by undergoing examination with the desired test battery prior to sustaining head injury. In military service in time of war it is safe to predict considerable numbers of brain injuries among casualties, but tests which can be *administered* economically to the enormous number of recruits in which these brain injuries later chance to occur are not likely to be sensitive to the subtle intellectual changes which are believed of clinical importance.

Huntington's chorea circumvents these sampling difficulties rather well in providing a population with a known incidence of easily recognizable and relatively homogeneous brain damage. It is a progressive degenerative disease of the central nervous system with an incidence (base rate) of about .0000543 in the general population (**14**). Thus, if one wished, he might draw a sample from the universe of infants and begin testing, secure in the knowledge that a known proportion of them would eventually develop Huntington's chorea and consequently be identifiable as definite cases of organic brain disease. Their rates of mental development and deterioration might be studied before and after symptoms developed, and might be contrasted with those for the remainder of the sample who did not develop this disease. Among the 8,000 infants already selected for our hypothetical longitudinal study of schizophrenia, the chances would be only a little better

than three in ten that even a single case of Huntington's chorea would appear. If we wished to ensure a working *N* of 68 cases of Huntington's, a sample drawn at random from the universe of infants would require an initial number something like 1,245,421—say a million and a quarter in round numbers.

Fortunately for investigative purposes, in Huntington's chorea we are not dependent upon random selection from the general population. Neither must we adhere to a strictly birth to death follow-up to ensure measures of premorbid and deteriorated levels of functioning in victims of the disease. By restricting our sample to children, one of whose parents was known to have Huntington's chorea, we immediately raise the base rate from .0000543 to .50000. This fact has been demonstrated empirically in repeated studies of choreic kinships which indicate that in Huntington's chorea the genetic concept of a single autosomal dominant gene with complete penetrance serves to predict and explain the observed facts.

Second, while the onset of symptoms in Huntington's chorea has been observed as early as birth and as late as age 77, the disease usually manifests itself in middle life. The average age of onset is about 36, with a standard deviation of around 12 years, and a distribution (allowing for probable systematic reportorial inaccuracies in observed data) which approximates a normal curve (**1**, **14**). Thus something like 68% of the cases will begin to show symptoms of Huntington's chorea while in the 24- to 48-year age range. If we were to begin with a sample of 100 persons of the same age selected at random from among all individuals with one choreic parent, and begin an intensive longitudinal study of those who appeared clinically normal at age 24, we might expect within 24 years to observe the onset of the deteriorating process in 34 cases. Thus, an *N* of 200 subjects selected on the basis of .50 genetic probability at birth would provide a working *N* of 68 clinical cases over a period of 24 years. Sixteen cases would be expected to appear before the study began and sixteen cases after it terminated.

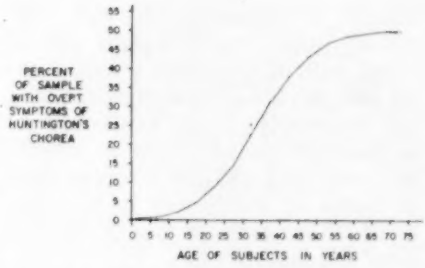
PERCENT OF SAMPLE WITH OVERT SYMPTOMS OF HUNTINGTON'S CHOREA
0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75
AGE OF SUBJECTS IN YEARS

FIG. 1. AGE-SPECIFIC EXPECTANCIES FOR HUNTINGTON'S CHOREA AMONG CHILDREN WITH ONE AFFECTED PARENT (BASED ON FINDINGS OF BELL AND PEARSON)

Figure 1 indicates the relationship between age and the probability of subsequently developing Huntington's chorea for children with one affected parent. It may also indicate to the reader appalled by the thought of a 24-year study that judicious selection of subjects in the age groups most liable to development of the disease would provide opportunity to observe its onset and course in a high percentage of cases within a much shorter period. In any case, however, human psychogenetic research requires a temporal reorientation for most behavioral scientists accustomed to projects of limited duration, quick results, and early publication.

Huntington's chorea is almost unique among mental disorders in that environmental variables appear to be of little significance in influencing the disease process. If the abnormal gene is present, the disease will manifest itself in any environment

capable of sustaining life, and individual differences as regards age of onset, rate of progression, neurological and psychological symptoms, etc., are seemingly more explicable in terms of different biotypes or genetic strains among afflicted kinships than in terms of differences in environment, either early or late in the victims' life histories (3).

In the more common "functional" mental disorders which confront thousands of psychological investigators and practitioners in their daily work, the situation is somewhat different. No one among contemporary authorities of genetic persuasion would deny that differences in environment contribute heavily to behavioral variance within and between abnormal groups. In spite of this fact, the extensive data on the familial incidence of mental disorders do afford the basis for developing age-specific expectancy charts with applications similar to those in our illustration with Huntington's chorea. To avoid the necessity for detailing the reasons psychologists give for generally ignoring geneticists' findings, we may assume that any tendency of mental disorders to run in families is attributable entirely to social inheritance. Without prejudice to the idea of preselected subpopulations, we may assume that the biological endowment of human embryos for potential tolerance to stress is substantially identical, subject only to differing experience during and subsequent to parturition.

Considering first schizophrenia, the example utilized at the beginning of this paper of a sample with an expectancy for the disease of 68% was not merely hypothetical. This figure is based upon the finding in pedigree studies by Kallmann (8), that 68.1% of children born to matings where both parents were schizophrenic developed the disease themselves. Where a single parent was known to be schizophrenic, 16.4% of the offspring developed the disease.

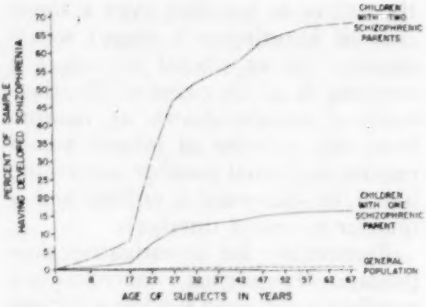
CHILDREN WITH TWO SCHIZOPHRENIC PARENTS
CHILDREN WITH ONE SCHIZOPHRENIC PARENT
GENERAL POPULATION
PERCENT OF SAMPLE HAVING DEVELOPED SCHIZOPHRENIA
AGE OF SUBJECTS IN YEARS

FIG. 2. AGE-SPECIFIC EXPECTANCIES FOR SCHIZOPHRENIA IN SPECIFIED POPULATIONS (BASED ON FINDINGS OF KALLMAN AND SLATER)

Determination of age of onset in schizophrenia presents somewhat greater problems than in Huntington's chorea, and reliable figures are in short supply. For purposes of psychogenetic investigation, a major complicating factor is the fact that early onset has been numbered among the diagnostic criteria for schizophrenia. Slater (20) has presented data on the ages at onset of schizophrenia among 67 siblings of 156 *propositi* or index cases, thus affording somewhat greater confidence that a reasonable picture of the incidence in older age groups is secured than would a much larger number of schizophrenic index cases diagnosed by conventional procedures in usual psychiatric practice.

Figure 2 reflects Slater's data on age of onset against Kallmann's risk figures for schizophrenia among subjects with one and with both parents known to be schizophrenic.

We must recognize possible limitations due to sampling and methodological errors in the studies cited. With this qualification, it is obvious that if an investigator seeks by a de-

velopmental approach to understand the crucial factors which distinguish the social inheritance of persons who become schizophrenic from that of their siblings and others who do not, he can spare himself considerable effort by selecting his cases from populations where probabilities are high that the necessary etiologic conditions will obtain.

The affective psychoses, likewise, have been shown to occur with greater frequency among relatives of affected *propositi* than among people in general. Since there is an ascertainable (though as yet inadequately investigated) pattern for age in relationship to frequency of diagnosis, it would be quite possible to develop age-specific expressions of probability for developing an affective psychosis based upon empiric risk figures already at hand for various subpopulations.

Manic-depressive psychosis appears to be somewhat less frequent in the general population (base rate about .004) than does schizophrenia. In spite of this fact, the economies effected by studying a subpopulation of persons, each with one manic-depressive parent, would be even greater than in the case of schizophrenia, in that about 23% of the offspring in matings with one manic-depressive parent will also develop the disease, as against the 16.4% morbidity in schizophrenia.

Geneticists have sought to explain these differences in base rates and familial morbidity (along with some of the other interesting findings in these disorders) in terms of a dominant mode of inheritance in manic-depressive psychosis versus a recessive in schizophrenia, with the introduction of several other constructs such as partial penetrance and polygenic modifying factors to account for some of the troublesome details. Persons emphasizing the factor of social inheritance might seek to explain these differences in other terms. Perhaps children simply find it easier to identify with a cyclothymic parent than with a schizoid one, the former being more likable in their better moments. Possibly because manic-depressive disorders are believed on the average to have a later onset than do the schizophrenias, children of the former have a longer average exposure time to the unfortunate parental influence, and consequently a heightened risk of patterning their personalities after undesirable models. Hypotheses could be multiplied, but the important thing for behavioral scientists is to examine the facts to see if they really do demand explanation. If so, the explanation should be sought empirically and experimentally where there is a reasonable prospect of finding it. This could well be in longitudinal studies of appropriate samples with high base rates.

Psychoneuroses and sociopathic personality disturbances are more difficult to study from the standpoint of inheritance than are the psychoses, and have received less attention from human geneticists, although the evidence from several sources is of interest (**6, 9, 21**). Traditionally, these conditions are regarded as purely psychogenic, even by the psychologists who admit the possibility of inherited biological predispositions toward psychotic modes of adjustment. Under our assumption concerning social inheritance, however, it seems reasonable and consistent with all current personality theories to predict that the incidence of neuroticism, psychopathy, etc., will be relatively high among the children of parents with these difficulties. Accordingly, such children should constitute an acceptable high base-rate

group for investigators with even the most rigid environmentalistic biases. The problem remains of determining base rates and age expectancies for these disorders. Unless this problem is overcome, longitudinal studies would be limited so far as assessment of mental hygiene measures is concerned. It should still be possible, however, to make important empirical observations of children of neurotic and psychopathic parents, and possibly to demonstrate the nature of the crucial differences between the psychological environments of those children who later become neurotic or psychopathic and the children who later remain well adjusted, despite sharing the same physical environments with their less fortunate sibs.

Presumably, base rates for frequency of neurosis and psychopathy in the general population would not run as high as the 19.04% judged neurotic and 3.56% judged psychopathic by Pearson and Amacher (**13**) in a large sample of unmarried mothers. Even if the over-all base rates were to approach or exceed these figures, the economies inherent in using samples selected because of predisposing social inheritance (and/or biological inheritance) should prove considerable.

In the foregoing examples, the acceptance of traditional broad diagnostic categories such as "the schizophrenias," "the affective psychoses," and "the psychoneuroses" may be deemed objectionable by some. The principle underlying a psychogenetic approach to behavior disorders remains the same, however, even if one were to restrict a study to the nature and frequency of neuroses in female children of mothers with conversion hysteria manifested by glove anesthesia of the left hand. The present need is for studies under broad rubrics, and the categories cited above are convenient starting points. Eventually, if the limits of prediction of behavior from knowledge of parental behavior can be demonstrated, a psychogenetic approach might lead to a highly refined classificatory system for behavior disorders, even if a crisp definition like that of "normal" and "affected" in Huntington's chorea cannot be attained elsewhere.

Previous Considerations of Genetic Factors in Psychological Research

Comparative psychologists have long recognized the need for control of biological inheritance in experiments with animal behavior and the possibilities for utilizing strains bred to provide high or low base rates for particular behaviors. "Maze-bright" and "maze-dull" rats produced through genetic selection are familiar to almost every student of introductory psychology, and even Huntington's chorea has its counterpart in the mouse (**23, 25**). The term "psychogenetics" appears first to have been used by Hall (**5**), who decried the relative lack of attention to inherited behavioral dispositions by psychologists, and who saw as a potent cause of individual differences within species, structures (genes) which can produce such varied forms as the flea and the elephant or the starfish and the kangaroo. Unfortunately, Hall foresaw little possibility for developing a science of human psychogenetics, apparently feeling that the long span of years per human generation, and the notorious reluctance of people to mate in accordance with an experimental design, were effective barriers to worthwhile application of genetic principles. Scott (**19**), likewise, saw genetics as a tool in psychological research, but only in terms of inbred and, therefore, subhuman strains.

As indicated earlier, psychological research involving genetic factors in human personality attributes to date has been geared almost exclusively to efforts to disclose the relative potency of heredity and environment in determining psychological characteristics. Extensive methodologies have been developed, and numerous efforts to apply these have been recorded, particularly as regards individual differences in human intelligence (11). Most authorities are in agreement that both heredity and environment have something to do with the production of such differences, but there has been a dearth of suggestions as to how this conclusion can serve as a steppingstone to better prediction and control of behavior. The fact that there is still room for wide disagreement between some who assign a minimal role to inheritance and others who feel that environment only operates to produce differences within narrow limits fixed by heredity is probably due primarily to the uncertain factorial composition of human intelligence, with consequent uncertainty as to the contribution of errors of measurement to total variance.

In studies of nature and nurture in relation to emotional adjustment, the difficulties in defining variables and in constructing valid and reliable measures have proved even greater than in the case of intelligence. The recent application of criterion analysis to measures of neuroticism among twins by Eysenck and Prell (4) and Cattell's suggestion (2) for use of the multiple-variance method in comparing populations with differing degrees of genetic and environmental homogeneity are interesting methodological advances. Both, however involve the difficulties generally applying in twin studies; evidence since Galton's early proposals has shown twinship to be considerably less than the ideal provision by nature for a planned experiment (16).

There is possibly an even more basic difficulty common to past efforts directed toward the nature-nurture problem in human behavior. This is in the proclivity of psychologists for assuming or demonstrating variables to be distributed in continuous fashion throughout the general population, but concentrating attention on the pathological extremes, which may, in fact, constitute discrete series. The familiar concept of a normal distribution for "emotional adjustment" ranging from "super normal" and "normal" to "neurotic" and "psychotic" has led laymen and behavioral scientists alike to picture human emotions in varying shades of gray. Various schemata have been proposed to place particular maladjustments on continua like that of Richards (18) who regards schizophrenia and hysteria as opposite poles of an extratension-introversion continuum. Still more specific kinds of behavior are treated as continuous. We may even regard the mother who projects the sensations from her own full bladder and insists that her young daughter has to go to the bathroom as different in degree, but not in kind, from a lady with the fixed false belief that the President has erotic designs upon her person.

While such conceptualizations may serve a useful purpose, they may also be misleading. The danger lies in the temptation to infer continuous distribution of underlying etiological factors from the fact that behavioral traits appear to be so distributed. Huntington's chorea may again illustrate the point: Symptomwise, victims of this disease might be regarded as the extreme of the distribution for motor steadiness or manual dexterity, which evidence suggests are normally

distributed in the general population. However, regardless of the nature of the causes of individual differences in these traits among people in general, it is clear that the Huntington's chorea victims constitute a group apart, different not in degree, but in kind, by reason of some specific biochemical error which is highly predictable in terms of inheritance, and which operates in a manner quite different from anything observed in nonchoreic kinships.

The study of Eysenck and Prell cited earlier led these authors to conclude that there is a biological factor of neuroticism in which biological inheritance contributes about 80% of the variance. However, their assumption that neuroticism is on a continuum in the general population and the samples employed make it impossible to infer that clinical cases of neurosis arise at the extreme of the continuum only because of the *degree* to which they inherit the neuroticism factor. Testwise or symptomwise the diagnosed neurotics do constitute the extreme of a distribution, but the reasons for their coming to this sorry end may be quite different from the reasons which cause individuals in the "borderline" or "normal" range of test scores or clinical behavior to fall where they do. Just as Huntington's choreics constitute a group apart (etiologically) on the motor-steadiness continuum, so clinical neurotics might fall at the end of the distribution by reason of a different *kind* of heredity, *either biological or social.* If one surveys the work on human inheritance, a general pattern emerges as regards etiological differences between normal and pathological degrees of variability. In genetic terms, normal variability of hereditary characteristics within a species is explained on the basis of polygenic inheritance. People who are very tall or very short, but still within the "normal" range, appear to manifest the influence of many genes or gene pairs, each with small but similar effect. At both extremes of the distribution for height, however, observations indicate that a number of hereditary pathological syndromes are determined by single genes or gene pairs of very potent influence in determining the metabolic processes of the entire organism. Chondrodystrophic dwarfism has been shown to be transmitted as a dominant in a number of kinships, as has its symptomatic inversion, arachnodactyly (9). At least some forms of gigantism operate as recessive characters (7).

With respect to psychological characteristics, some direct evidence is provided in the area of intelligence. Again, from the genetic point of view, studies showing high parent-child correlations on measures of intelligence throughout the normal range are interpreted as indicating that numerous gene pairs are involved in the production of these individual differences. At the low extreme of the distribution, however, the existence of several etiological agents transmitted after the fashion of single genes or gene pairs has been demonstrated. Among congenital varieties of mental deficiency, microcephaly, Tay-Sachs disease, or infantile amaurotic idiocy, and several other syndromes have appeared as recessive characteristics in numerous kinships, while others such as epiloia behave as dominants. Dementing conditions in which individuals seemingly normal for a time after parturition deteriorate intellectually and develop the appearance of mental deficiency weeks, months, or years following birth likewise may show either dominant or recessive inheritance. Phenylpyruvic acid oligophrenia, juvenile amaurotic idiocy,

Alzheimer's disease, and, of course, Huntington's chorea are prime examples (15). The writers are even of the opinion that at the extreme upper end of the distribution for intelligence, as well, single-factor biological inheritance may possibly account for some of the gross anomalies observed from time to time.

The inference that the pathological extremes of the "emotional adjustment continuum" may be determined by unitary factors of biological inheritance because this seems to be the case with other physical and mental traits of humans is provocative. However, as in the case of other analogies, it may not be regarded as proof. To do so would be an error of the same sort as is commonly committed in inferring etiological factors for maladjustment to be distributed along a continuum from the fact that symptomatic behaviors appear to be so distributed. However, when the weight of evidence from direct genetic studies of emotional disorder is thrown into the balance along with the suggestive analogy above, the possibility appears very real that the individuals who become ill with any of several "functional" disorders are, in fact, discrete populations, biologically predisposed by unitary inherited factors toward particular kinds of abnormal behavior now generally regarded as arising from an unfortunate kind or number of learning experiences.

The writers' bias in emphasis on biological inheritance is by now apparent. The fact remains, however, as stated at the outset, that the longitudinal study of specified subpopulations is equally consonant with an environmental bias, to say nothing of an open-minded approach. The base-rates problem in regard to etiology of mental illness is a very real and present one for psychological science. So is the possibility that the presence of the single anomalous molecule of protein which we term an abnormal gene may override all other considerations in the prediction of abnormal behavior. If the mainstream of psychological thought continues to ignore these problems, the result might be the loss of an opportunity to foreshorten by years the search for truly effective means of preventing and treating mental illness. Beyond this, if these problems are rejected as not within our province, the result might be the abdication by psychology of its cherished role as a leader in objectively oriented research among the behavioral sciences.

Practical Problems in Human Psychogenetic Research

In the foregoing section, the suggestion is implicit that mental disorder is in large part a genetic and, therefore, ultimately a biochemical problem. Assuming this to be true, the question immediately arises, "Why not leave it to the geneticists and biochemists to work it out, and let behavioral scientists struggle along with present methods for dealing with these real-life problems until something better can be devised?" The answers to this question are several. Probably foremost is the fact that scientists have or should have an interest in being right, as well as in being effective. A given theory may do fairly well in subsuming observations and enabling predictions, but it is the mark of the scientist not to rest until a better system is devised. Certainly the present status of prediction and control of human behavior disorders allows little room for satisfaction among its serious students, let alone smugness.

A second reason why psychology should concern itself with psychogenetics in relation to mental disorder

lies in the relative numbers and locations of persons with requisite training in either psychological or gènetic research methodology. Persons with graduate-level training in human genetics are as yet few in number, and tend to find employment almost exclusively in university settings. Psychologists, on the other hand, are relatively numerous and abound in clinics and hospitals where the index cases necessary to psychogenetic investigation present themselves.

A third factor favoring hybridization of psychological and genetic sciences in this area is the positive contribution which psychology can make in speeding the progress of work from a strictly genetic point of view. Procedures for establishing and maintaining rapport with human subjects and principles of communication with persons lacking in intelligence and not familiar with scientific jargon are two of several areas in which geneticists might profit from the experience of psychologists.

Granting that increased cross fertilization between psychology and genetics is desirable in the conception of research on mental disorders, there remain several questions as to the practicality of longitudinal investigations as proposed earlier. Realistically, what are the chances that a population of infants could be identified as each having two schizophrenic parents? The chances are, perhaps, much better than they appear at first glance. If investigators in various locations were alerted to the potential importance of such subjects in subsequent research, even the seemingly irreducible incidence of matings by schizophrenics within the best-run mental hospitals across the country in a single year would probably provide a sample large enough to be reckoned with when one considers the base rate or expectancy of 68%. Subjects selected on this basis would, almost without exception, be placed in adoptive homes. They could be compared with other adopted children of nonschizophrenic parents placed by the same agencies, to give an indication of the importance of adoptive placement per se in the etiology of schizophrenia or such other disorders as might appear in the two groups.

The morbidity expectancies for other subpopulations discussed earlier would be lower, but, by the same token, the subjects would be more readily available. Children with one normal and one schizophrenic or manic-depressive parent would present no particular problems. Recognized matings of two psychopaths (again with a high proportion of offspring placed in adoptive homes) are not uncommon, and there is quite possibly an affinity of male neurotics for female neurotics, which could be turned to account. In any case, even if much more restrictive diagnostic categories were employed, the incidence of usable cases would probably be many times greater than that of monozygotic twins reared in dissimilar environments—a kind of subject which, by virtue of its rarity, has proved the stumbling block in many well-conceived experiments on genetic factors in human behavior.

Given the names and addresses of a sizable list of persons selected for a particular psychogenetic study, what would be the chances of enlisting the cooperation of all the interested parties necessary for repeated submission to psychometric and other studies over a period of years? Again, the possibilities are probably not as remote as they may appear. Terman's studies of genius have shown that subjects in longitudinal studies will remain in contact with an investiga-

tor for years, expend a great deal of effort, and divulge all sorts of highly personal information with no reward other than the satisfaction of being singled out for study as intellectually gifted. Our own work in Huntington's chorea indicates that a large majority of persons with 50% probability for developing the disease are willing to participate in efforts to predict and control the disease. Over 200 subjects have evidenced this willingness in submitting to psychological and electroencephalographic examination on one or more occasions. Many of them have traveled hundreds of miles at their own expense to do so, despite foreknowledge that the results of the study were likely to be of no direct or immediate benefit to them or their families.

This experience may have no implications for similar approaches to "functional" disorders where the familial morbidity is lower and where the exact nature of the etiology is in doubt. However, this is no basis for rejecting the approach without trying it. A number of authorities questioned on a priori grounds the feasibility of informing large numbers of choreic families about the inheritance of the disease and of inviting their participation in predictive studies only to have the fears proved groundless when an effort was made. The same finding might very well obtain in populations selected to show a high incidence of schizophrenia or other disorders if a frank and simple presentation were made of the nature and purpose of the study.

To judge from our experience with persons under the dire threat of Huntington's chorea, the picture of human personality (even the personality one might expect to encounter among a group of preschizophrenic or preneurotic individuals) as a fragile thing likely to be precipitated into catastrophic reaction by the intimation that emotional disorders may run in one's family, is a specter which exists in the minds of behavioral scientists and not, as a rule, among the persons most directly concerned. Instead, one might expect to find a certain toughness of psyche and a desire to face facts, coupled with the feeling that such authorities in the field of mental health as they may have consulted have not been completely (*a*) informed or (*b*) honest concerning the etiology of emotional disturbances.

The recent rise of quiz shows, opinion polls, and surveys, together with the generally heightened public awareness of the importance of scientific research, appear to have created an atmosphere conducive to cooperation in any investigation under legitimate sponsorship, even if it involves the sort of data collected in the Kinsey studies. Asking laymen to take voluntarily a series of personality tests, or to have their brain waves recorded for the sake of science, is now likely to evoke general acceptance of a sort rare only a few years ago except among college sophomores. The limits of this public warming toward scientific investigation have not been exploited, nor, perhaps, should they be. The fact remains, however, that one cannot assess the reaction of the population involved toward a psychogenetic research design until he tests it. In Huntington's chorea, which in certain respects constitutes an acid test, the reactions have been favorable.

Possibilities also exist for utilizing a psychogenetic approach to the study of human emotions without ever divulging to the subjects the fact that they were special objects of study. In Cattell's efforts to apply

the multiple-variance method, for example, it was proposed to test the entire school grades in which selected subjects such as twins reared by different foster parents chanced to be placed.[1] Thus children, parents, and school authorities would know at most that a particular classroom was under consideration, without further identification of the desired subjects except by the investigator.

In short, the practical obstacles to psychogenetic research in terms of availability of subjects do not appear insurmountable. As indicated earlier, patience and forbearance are prerequisite virtues in investigators seeking to use this approach. Financial support is not easily obtained for proposals which reckon in decades rather than in months the interval before results will be forthcoming. Again, however, such problems should not prove insurmountable. Probably the immediate need is for additional studies which would verify or modify the base-rate expectancies and ages of onset for different disorders in different subpopulations. These could be completed in a relatively short period if any considerable number of psychologists or other behavioral scientists were persuaded to undertake them.

In the longitudinal study of samples drawn from such subpopulations, cooperative effort by a number of investigators, each studying a few individuals spread across the country might prove a successful alternative to a study conducted by a single research center. It is obvious that a multidisciplinary approach should prove most fruitful in such research. For example, if someone went to the trouble of making arrangements necessary for the study of 100 infants with 68% expectancy for schizophrenia, the experimental design should not be restricted to intelligence and personality testing at specified intervals. Rather, the best available talent in anthropology, sociology, biochemistry, neurophysiology, biophysics, and other disciplines should be focused on such a unique opportunity. The question as to who will take the lead in creating such opportunities remains open.

[1] Cattell, R. B. Personal communication, 1954.

Summary

The long recognized tendency of many abnormal behavior patterns to run in families may prove useful in understanding, predicting, and controlling such behavior, regardless of whether the ultimate cause lies in social heredity or in biological heredity, and regardless of the extent to which these are correlated. Studies by human geneticists and others provide data which make it possible to calculate age-specific statements of probability for developing various disorders in various subpopulations defined in terms of genetic relationship to index cases. These subpopulations might be studied longitudinally in efforts to disclose the crucial differences in heredity (either biological or social) between subjects who develop the disorders in question and those who do not. Studies of such subpopulations would probably prove much more economical than would investigations in random samples from the general population, by reason of the much higher base rates for incidence of the specified disease in the former. Preventive or mental hygiene measures could also be evaluated economically in longitudinal efforts with subpopulations having known base rates. The base-rates problem is discussed with respect to predicting psychiatric illness from kind and number of

traumatic life experiences, and it is suggested that this problem has been grossly neglected in behavioral science. Examples of genetic expectancies as age-specific base rates for Huntington's chorea and schizophrenia are presented and the possibilities for similar expressions in other disorders are discussed. Much previous psychological research on the nature-nurture question may involve the erroneous assumption that because behavioral variables appear to be continuously distributed in the general population, the underlying etiological factors are also thus distributed. Examples are cited to show that from the etiological standpoint the extremes of distributions for several physiological and psychological characteristics actually constitute discrete series. Practical problems with respect to psychogenetic research on human behavior disorders are discussed. It is recommended that psychologists and others interested in human behavior give further consideration to possibilities in this area.

REFERENCES

1. BELL, JULIA. Nervous diseases and muscular dystrophies: Huntington's chorea. Vol. 4, Pt. 1, *Treasury of human inheritance.* Cambridge: Cambridge Univer. Press, 1935. Pp. 1–68.
2. CATTELL, R. Research designs in psychological genetics with special reference to the multiple variance method. *Amer. J. hum. Genet.*, 1953, **5**, 76–93.
3. DAVENPORT, C., & MUNCEY, ELIZABETH. Huntington's chorea in relation to heredity and eugenics. *Amer. J. Insan.*, 1916, **73**, 195–222.
4. EYSENCK, H., & PRELL, D. The inheritance of neuroticism: an experimental study. *J. ment. Sci.*, 1951, **97**, 441–465.
5. HALL, C. The genetics of behavior. In S. S. Stevens (Ed.) *Handbook of experimental psychology.* New York: Wiley, 1951.
6. JOST, H., & SONTAG, L. The genetic factor in autonomic nervous system function. *Psychosom. Med.*, 1944, **6**, 308–310.
7. KALLMANN, F. *Heredity in health and mental disorder.* New York: W. W. Norton, 1953.
8. KALLMANN, F. The genetic theory of schizophrenia: An analysis of 691 schizophrenic twin index families. *Amer. J. Psychiat.*, 1946, **103**, 309–322.
9. KEMP, T. *Genetics and disease.* Copenhagen: Ejnar Munksgaard, 1951.
10. MEEHL, P. E., & ROSEN, A. Antecedent probability and the efficiency of psychometric signs, patterns, or cutting scores. *Psychol. Bull.*, 1955, **52**, 194–216.
11. NATIONAL SOCIETY FOR THE STUDY OF EDUCATION. *The thirty-ninth yearbook.* Bloomington, Illinois: Public School Publishing, 1940.
12. OLTMAN, J., MCGARRY, J., & FRIEDMAN, S. Parental deprivation and the "broken home" in dementia praecox and other mental disorders. *Amer. J. Psychiat.*, 1952, **108**, 685–694.
13. PEARSON, J., & AMACHER, PHYLLIS. Intelligence test results and observations of personality disorder among 3594 unwed mothers in Minnesota. *J. Clin. Psychol.*, 1956, **12**, 16–21.
14. PEARSON, J., PETERSEN, M., LAZARTE, J., BLODGETT, HARRIET, & KLEY, IRENE. An educational approach to the social problem of Huntington's chorea. *Proc. Staff Meetings Mayo Clinic*, 1955, **30**, 349–357.
15. PENROSE, L. *The biology of mental defect.* (2nd ed.) London: Sidgwick & Jackson, 1954.
16. PRICE, B. Primary biases in twin studies. *Amer. J. hum. Genet.*, 1950, **2**, 293–352.
17. RENAUD, H., & ESTESS, F. Life history interviews with 100 normal American males: Pathogenicity of childhood. *Amer. Psychologist*, 1955, **10**, 371.
18. RICHARDS, T. *Modern clinical psychology.* New York: McGraw-Hill, 1946.
19. SCOTT, J. Genetics as a tool in experimental psychological research. *Amer. Psychologist*, 1949, **4**, 526–530.
20. SLATER, E. Psychotic and neurotic illnesses in twins. *Med. Res. Council: Spec. Rep. Ser.*, 1953, **278**, 1–358.

21. SLATER, E. Psychopathic personality as a genetical concept. *J. ment. Sci.*, 1948, **94**, 277–282.
22. SLATER, E., & SLATER, P. A heuristic theory of neurosis. *J. Neurol. Neurosurg. Psychiat.*, 1944, **7**, 49–55.
23. TRYON, R. Individual differences. In F. A. Moss (Ed.) *Comparative psychology*. New York: Prentice-Hall, 1942.
24. YATES, A. The validity of some psychological tests of brain damage. *Psychol. Bull.*, 1954, **51**, 359–379.
25. YERKES, R. *The dancing mouse*. The Animal Behavior Series. No. 1. New York: Macmillan, 1907.

Received May 20, 1957.

PSYCHOLOGICAL BULLETIN
Vol. 54, No. 5, 1957

THE -ILES THAT PLAGUE ELEMENTARY STATISTICS

HORACE B. ENGLISH
Ohio State University

Every instructor who attempts to instill a few notions of elementary statistics in his students knows how hard it is to gain acceptance for the designation of the top or best as the fourth quartile. After all, the student has been accustomed most of his life to speaking of the best quarter of his class as the *first*. "A quartile *is* just a quarter, isn't it?"

It is unfortunate that the statisticians who initiated the numerical designation for quartiles and other partiles did not think of possible confusion with a common speech usage; but the ranking of all the partiles from the bottom up is now too well established in technical usage to be changed. Rationalization of the names and numerical designations of the several concepts involved in dividing a ranked distribution is, however, highly desirable in the interests of easy communication. In working through the tangle for the *Dictionary of Psychological and Psychoanalytic Terms*,[1] I developed a schema that is relatively simple and self-consistent, and that puts a minimum strain on established technical usage or on ordinary habits of thought—the last being a most important criterion for good terminology.

A *partile* is defined as "the generic name for one of the set of points that divide a serially ordered or ranked distribution into a number of divisions, each of which contains the same number of scores. Each point is located as coinciding with the obtained score in the distribution below which the required fraction of the obtained scores is found. If there are 100 such divisions (each containing 1/100 of the scores), the division points are called *percentiles* (sometimes *centiles*, but this usage, though logical, conflicts with another established usage). Other partile points are named by adding *-ile* to the root of the Latin ordinal number: e.g., *decile* (setting off divisions of 1/10 of the cases), *octile* (1/8), *sextile* (1/6), *quintile* (1/5), *quartile* (1/4), *tertile* (1/3)."

Note that partile is defined as a point. This has always been the usage of careful writers, but the introduction of *percentile* for the more logical *centile* to designate the dividing *point* left centile with no job. It quickly took over the task of naming a hundredth part of a ranked distribution; and this sort of designation, extended downward, led to naming the tenth part a decile, the fourth part a quartile, etc.

The lower ranking *-ile* terms—tertiles, deciles, etc.—were thus burdened with a double task: to name the dividing point but also the parts separated by the points. This does not seem too confusing until we note that, in the case of the quartile, e.g., the top quartile *division* is the fourth while the top quartile *point* is the third. A good mathematical mind says "Of course" but the average student says "How come?" Move up, moreover, to the higher partiles and ask even a good statistical student where the 95th centile division is—whether it is fourth, fifth, or sixth from the top. He won't be able to answer quickly; and errors in designation, resulting in misrepresentation of the facts, are not uncommon.

It won't avoid all confusion but it

[1] Longmans Green, New York, 1958.

will minimize it if we follow best usage and restrict the *-iles* when written without qualifiers to the *dividing points*. (The only exception to this rule is centile and even that might well be avoided as shown below.)

How then shall we name the divisions or parts set off by the partile points? My own preference is simple. What are the two parts divided by the median? Even in the most technical tomes, they are halves. Why not then thirds, fourths, sixths, tenths, hundredths? If the context makes clear that we are dealing with a ranked distribution, these are wholly unambiguous. But for those who prefer a term that looks a little more technical, *tertile division*, *decile division*, etc., are also unambiguous. But just plain decile is not. If we call the division a hundredth, we get rid of the illogicality of centile. If we must use an *-ile* term, we should parallel *tertile division* with *centile division*, not just centile.

A third concept is involved: the range of scores in any division. This has been called the *partile interval* (quartile, quintile, etc., interval) and is clear.

There remains finally the question of ranks. What about such terms as decile rank, percentile (or is it centile) rank? Of course, the designating number given to either the dividing point or to the division is its rank. The rank of quartile three (the point divider) is 3. The rank of the top percentile is 99.

But rank in this connection is commonly used to assign a rank order to the *individual:* to a score, a person, a particular item. And this is done by assigning to this individual the rank of the *partile division* in which the individual is found. Thus the rank of an individual is accurately given as *quartile division rank*, and as *centile division rank* (*not* percentile rank, which introduces not only semantic but numerical error). This may be expressed elliptically as *quartile rank* or *centile rank* but we must remember it is the rank of the division, not of the dividing point.

All of this will seem to the working statistician to belabor the obvious. It is not the obvious for the beginning student. Nor is it simple. It is not as simple a scheme as one could devise if we were beginning afresh. A scientific terminology usually grows by accretion, often in peculiar directions. But this scheme seems to put as much clarity and rationality into the pattern as current usage permits.

Received April 8, 1957.

PSYCHOLOGICAL BULLETIN
Vol. 54, No. 5, 1957

A NOTE CONCERNING KENDALL'S TAU[1]

DESMOND S. CARTWRIGHT
University of Chicago

In a recent article in this journal (**4**), Schaeffer and Levitt have provided a valuable survey of procedures for Kendall's τ-coefficient of rank-correlation (**3**). One feature of their presentation, however, may cause readers to employ significance tests yielding spurious results.

They define the τ-coefficient thus (**4**, p. 339):

$$\tau = \frac{S}{n(n-1)/2}. \qquad [1]$$

They state that when there are ties Eq. 1 may also be used when agreement with an untied objective ranking is being determined (**4**, p. 340). They continue (**4**, p. 340): "In general, however, the corrected formula should be used since rank correlations are usually computed when agreement rather than accuracy is the issue." For the corrected formula they specify (**4**, p. 339)

$$\tau = \frac{S}{\sqrt{n(n-1)/2 - V} \cdot \sqrt{n(n-1)/2 - U}}, \qquad [2]$$

where

$$V = \frac{1}{2}\sum_v v(v-1)$$

and

$$U = \frac{1}{2}\sum_u u(u-1).$$

In their footnote number 5 (**4**, p. 341) the authors present the basic equation and several derivative equations for the variance of τ when there are ties. They state that these formulas are to be found in Kendall (**3**). In fact, they do not appear in the same form in Kendall's discussion. The latter gives expressions for the variance of S, the numerator of Eq.1 and Eq. 2.

Kendall's basic equation is (**3**, pp. 43, 57–60)

$$var\ S = \frac{1}{18}\left(n(n-1)(2n+5) - \sum_v v(v-1)(2v+5) - \sum_u u(u-1)(2u+5)\right)$$
$$+ \frac{1}{9n(n-1)(n-2)}\left(\sum_v v(v-1)(v-2)\right)\left(\sum_u u(u-1)(u-2)\right)$$
$$+ \frac{1}{2n(n-1)}\left(\sum_v v(v-1)\right)\left(\sum_u u(u-1)\right). \qquad [3]$$

The basic equation and the derivative equations for the variance of τ as given by Schaeffer and Levitt (**4**, p. 341) may be expressed as

$$\sigma_\tau^2 = \frac{4}{n^2(n-1)^2} \cdot var\ S, \qquad [4]$$

where *var* S is given by Eq. 3. The first term on the right of Eq. 4 is equal to the reciprocal of the denominator of Eq. 1 squared. Now, letting D denote the denominator of

[1] This note was prepared in connection with research supported in part by funds from the Ford Foundation (Psychotherapy Research Fund), and in part by a grant (PHS M 903) from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service.

a formula for τ when there are ties, we may write

$$\sigma_\tau^2 = var\ S/D^2, \qquad [5]$$

which is true for all formulas, given only that D is in fact the denominator for the particular τ whose variance is to be computed. But Eq. 4, and all variance expressions given by Schaeffer and Levitt (4, p. 341) are applicable only to Eq. 1, not to the general formula for ties Eq. 2.

Moreover, in discussing corrections for continuity on S (C_s) when there are ties (4, pp. 341–342), the authors recommend that C_s be divided by $n(n-1)/2$ in order to obtain the correction for τ.

Now, we may write

$$C_\tau = C_s/D, \qquad [6]$$

which is true for all formulas for τ when there are ties, given only that D is in fact the denominator for the particular τ for which C_τ is to be computed. But the division recommended by Schaeffer and Levitt is applicable only to Eq. 1, not to the general formula Eq. 2.

Let us now examine a case illustrating the kind of error that may result from failure to apply Eq. 5 and Eq. 6 correctly. Eighteen boys and 16 girls are ranked on composite letter grades over a series of essays. The general formula (Eq. 2) is used to measure the association, if any, between sex and grades.

We have:

$$\left.\begin{aligned} n &= 34 \\ S &= 106 \\ V &= 273 \\ U &= 59 \\ D &= 380 \cdot 235 \end{aligned}\right\} \quad \tau = \frac{S}{D} = +.279.$$

We also find

$$var\ S = 3261.7754,$$

$$\sigma_s = \quad 57.11,$$

$$C_s = (68\text{–}1\text{–}1)/26 = 2.54,$$

[Whitfield's correction (5),]

$$S \text{ corrected} = S_c = 106\text{–}2.54 = 103.46.$$

Then, using the normal probability integral (3, p. 143) we find

$$\frac{S_c}{\sigma_s} = \frac{103.46}{57.11}\sigma = 1.81\sigma; \quad p < .08.$$

If now we employ the correction and variance expressions given by Schaeffer and Levitt, we find

$$\sigma_\tau^2 = .010364,$$

$$\sigma_\tau = .1018,$$

$$c_\tau = 5.08/1122 = .0045,$$

$$\tau \text{ corrected} = \tau_c = .279 - .0045 = .2745,$$

and

$$\frac{\tau_c}{\sigma_\tau} = \frac{.2745}{.1018}\sigma = 2.70\sigma; \cdot\ p < .008.$$

The discrepancy between $p < .08$ for S_c and $p < .008$ for τ_c is due solely to our failure to apply Eq. 5 and Eq. 6 correctly. Our results fail to satisfy the identity

$$\frac{S_c}{\sigma_s} \equiv \frac{\tau_c}{\sigma_\tau}, \qquad [7]$$

which must hold if it is accepted that a given result has one and only one p value.

Now, if we use Eq. 1 to determine τ for this example, and continue to use the correction and variance expressions given by Schaeffer and Levitt, we find

$$\tau = .1889,$$

$$\tau_c = .1844,$$

and

$$\frac{\tau_c}{\sigma_\tau} = \frac{.1844}{.1018}\sigma = 1.81\sigma; \quad p < .08.$$

This result satisfies Eq. 7.

Alternatively, if we use Eq. 2 as before but compute the variance of τ by Eq. 5 and the correction by Eq. 6, we have

$$\tau = .279,$$

$$\sigma_\tau^2 = .02256,$$

$$\sigma_\tau = .1502,$$

$$C_\tau = .0067,$$

$$\tau_c = .279 - .0067 = .2723,$$

and

$$\frac{\tau_c}{\sigma_\tau} = \frac{.2723}{.1502}\sigma = 1.81\sigma; \quad p < .08.$$

This result also satisfies Eq. 7.

As Schaeffer and Levitt point out, the formulas for ties in τ are complicated; but not more so than those for rho when the latter is properly employed. The authors discuss several advantages of τ over rho. A serious disadvantage has been the computational labor in obtaining S, especially when n is large. However, reasonably rapid procedures have now been developed by Bright (**1**) and by Cartwright (**2**). In the latter article it is also pointed out that it is not necessary to transform raw scores to ranks when computing S, which provides a further advantage over rho.

In testing the significance of an obtained τ when there are ties, it is probably best to follow the routine of first computing *var* S by Eq. 3 or by the derivatives of Eq. 3 for different conditions of ties as given by Kendall (**3**, pp. 43–45). The quantity C_τ may then be computed by the methods described by Schaeffer and Levitt (**4**, pp. 341–342). Using the appropriate D, σ_τ^2, and C_τ may then be computed by Eq. 5 and Eq. 6, respectively.

REFERENCES

1. BRIGHT, H. F. A method for computing the Kendall *tau* coefficient. *Educ. psychol. Measmt*, 1954, **14**, 700–704.
2. CARTWRIGHT, D. S. A computational procedure for *tau* correlation. *Psychometrika*, 1957, **22**, 97–104.
3. KENDALL, M. G. *Rank correlation methods*. London: Griffin, 1948.
4. SCHAEFFER, M. S., & LEVITT, E. E. Concerning Kendall's *tau*, a nonparametric correlation coefficient. *Psychol. Bull.*, 1956, **53**, 338–346.
5. WHITFIELD, J. W. Rank correlation between two variables, one of which is ranked, the other dichotomous. *Biometrika*, 1947, **34**, 292–296.

Received April 1, 1957.

PSYCHOLOGICAL BULLETIN
Vol. 54, No. 5, 1957

COMMENT ON A DISTRIBUTION-FREE FACTORIAL-DESIGN ANALYSIS

FRED D. SHEFFIELD
Yale University

In a recent article, Wilson (**4**) has proposed a distribution-free method for analyzing the results of factorial-design experiments. The ideal case involves running an equal number of subjects in each cell and dichotomizing at the over-all median of the quantitative scores obtained. The values analyzed are the *frequencies* above and below the over-all median. Subsequently Alluisi (**1**) has presented computational formulas for Wilson's method.

Perhaps because the revised cell values are frequencies, Wilson and Alluisi used a χ^2 approach in their analysis. With this approach they adduce a new set of formulas for coping with factorial design in terms of χ^2. They could, however, have treated the frequencies as scores and used the already well-known procedures of analysis of variance. The main purpose of the present comment is to reinterpret Wilson's method in more familiar terms. This purpose seems salutary if only to simplify the problems of the teacher and student of statistics, a field which has become overly complex and specialized in recent years. If a student knows analysis of variance he can handle this special case without recourse to Wilson's elaboration of χ^2 formulas.

In the ideal case of the Wilson method, the null hypothesis is that each observation in a cell has a 50-50 chance of falling above the over-all median. If N is the number of observations per cell, then the range of possible frequencies above the median is from 0 to N and the mean is $N/2$. As is well known, the variance of a frequency is NPQ, which in the present instance is $N(.5)(.5) = N/4$, since the null hypothesis is that $P = Q = .5$ cases above the median. This value of $N/4$ is therefore the within-cell variance. Moreover, it is a theoretically given within-cell variance, as if the degrees of freedom for its "estimation" were infinite.

From this point on the analysis of variance of the frequency scores can proceed in the usual way. This is exemplified below with the illustrative data used by Wilson. These data involve a 3×3 factorial design with 16 cases per cell. Thus the range of cell scores is from 0 to 16, the mean cell score is 8, and the variance of cell scores on the null hypothesis is 16(1/4), or 4. The obtained cell scores in Wilson's example are shown in Table 1. It should be understood

TABLE 1

		Illumination			
		1	2	3	Total
Dials	A	14	12	11	37
	B	9	7	8	24
	C	6	3	2	11
	Total	29	22	21	72

that while these cell values are frequencies above the median they are nevertheless scores in the same sense that the yield of a plot is a score in the agricultural experiments used as illustrations by R. A. Fisher (**3**). They differ in being less continuous, but are little more "rounded" than if one had coded a continuous distribution into about 17 class intervals.

Most students familiar with orthodox analysis of variance will be able to follow the analysis of Wilson's data shown below. The analysis should be regarded as one in which there is only one score per cell but in which there is an a priori within-cell variance. The computations below work from marginal totals rather than marginal means, but this is the usual approach in most texts which show computational methods for analysis of variance (e.g., **2**).

Total Σ Squares:

$$=(196+144+121+81+49+64+36+9+4)-\frac{(72)^2}{9}$$

$$=128$$

Σ Squares for Dials:

$$=\frac{[(37)^2+(24)^2+(11)^2]}{3}-\frac{(72)^2}{9}$$

$$=112.67$$

Σ Squares for Illumination:

$$=\frac{[(29)^2+(22)^2+(21)^2]}{3}-\frac{(72)^2}{9}$$

$$=12.67$$

Σ Squares for Interaction:

$$=128-112.67-12.67$$

$$=2.67$$

Since the error term is 4, the above Σ squares give rise to the analysis of variance in Table 2. The implications of these F tests are exactly the same as those of Wilson's χ^2 analysis. In fact, since χ^2 divided by its df is distributed the same as F for infinite df in the smaller variance, the present F values can be transformed into Wilson's χ^2 values by multiplying F by df (except for some minor calculation errors in Wilson's paper).

Thus far the present comments have been oriented toward a translation of Wilson's method into conventional analysis of variance. From this standpoint the above table is as far as analysis of variance can be carried with the nonparametric assumption that the within-cell variance is $NPQ=4$. Whether the above F tests are made or whether Wilson's χ^2 analysis is made, a single theoretical variance is the only error term possible, and all marginals and interactions must be tested against this single (within-cell) error term. Thus it is not possible to compare the mean squares of the two marginals with each other or with the Interaction mean square since these latter are empirical, i.e., parametric, values. This represents a fairly important limitation on what can be done with the nonparametric approach.

The importance of this limitation is apparent in the above data, in which F for Illumination is not at all significant by the nonparametric test but would be well within the 5 per cent level if tested in the conventional

TABLE 2

	Squares	*df*	Mean Square	*F*	*p*
Dials	112.67	2	56.34	14.08	$<.01$
Illumination	12.67	2	6.34	1.58	$>.05$
Interaction	2.67	4	0.67	0.17	—
Total (between cells)	128.00	8	16.00	4.00	$<.01$

way. Thus in a typical 3×3 factorial design experiment with only one score per cell there is no within-cell error because of the lack of replications. The only error term available in such a case is the Interaction of the two marginal variables. If this parametric approach is applied above, the F for Illumination against interaction is 6.34/0.67, or about 9.5, which is well beyond the 6.94 needed at the 5 per cent level for 2 and 4 df, respectively. The corresponding nonparametric test ($F = 1.58$) does not even reach the 20 per cent level of confidence.

The reason for this gross discrepancy in p values is readily apparent. The over-all F by the nonparametric test immediately disproves the null hypothesis and makes it not only an insensitive but also an untenable assumption against which to test for possible effects of a relatively weak marginal. The effects of Dials happens to be a strong one which holds up for all three Illumination settings and is significant even with the nonparametric test. This implies that the Illumination groups are *stratified* on an important variable, and while the Illumination means (or totals) are fairly close together, they differ more than chance allows for samples stratified on such an important variable. Another way of saying the same thing is that whereas the Illumination effects are small, they tend to hold up under all Dials conditions. Hence the significant F for Illumination against Interaction. The import is that the nonparametric approach will detect large effects but is relatively insensitive to small effects which are demonstrated by the data but which are not detected by the false null hypothesis.

In concluding this comment it should be noted that the Wilson method actually involves two separate parts. One part is a procedure for creating an approximately normal set of scores from an originally nonnormal set. This part of the method consists of dichotomizing at the median (or some equally reasonable cutting point) and making a score out of the number of values falling above this cutting point. The second part is a procedure for testing obtained variances against the theoretical variance implied by the cutting point used (i.e., $\sigma^2 = NPQ$). Only the second part of the method is the distribution-free part, and it suffers from the limitation that if the nonparametric hypothesis is proven false for the data as a whole, there is no proper method of testing the individual factors without recourse to parametric methods. A possible compromise is to use the nonparametric error term to test over-all F but to resort to parametric error terms if the over-all F disproves the original null hypothesis. This compromise is not possible within the framework of Wilson's χ^2 approach, which permits only comparisons between empirical variances and the theoretical variance. It is naturally suggested, however, by the present reinterpretation of Wilson's method in terms of conventional analysis of variance.

REFERENCES

1. ALLUISI, E. A. Computational formulae for a distribution-free test of analysis-of-variance hypotheses. *WADC Technical Report 56-339*, 1956.
2. EDWARDS, A. L. *Statistical methods for the behavioral sciences.* New York: Rinehart, 1954. Pp. xvii, 542.
3. FISHER, R. A. *The design of experiments.* London: Oliver & Boyd, 1947. Pp. xi, 240.
4. WILSON, K. V. A distribution-free test of analysis of variance hypotheses. *Psychol. Bull.*, 1956, **53**, 96–101.

Received May 6, 1957.

PSYCHOLOGICAL BULLETIN
Vol. 54, No. 5, 1957

THE STATISTICAL CONCEPTS OF CONFIDENCE AND SIGNIFICANCE

ROBERT E. CHANDLER
General Motors Employe Research

Recently there have been at least three different book reviewers who have commented on the confusion that currently exists in the psychological literature regarding the statistical concepts of confidence and significance (**2**, **7**, **9**). Although this confusion can be partially explained as a semantic problem, it behooves the psychologist to examine these two concepts rather closely and to adopt pristine terminology for the benefit of beginning students and individuals of other disciplines that draw rather heavily upon the psychological literature.

CONFIDENCE AND CONFIDENCE COEFFICIENTS

Confidence, a concept customarily reserved for discussions of interval estimation, is the faith which one is willing to place in a statement that an interval established by a sampling process actually contains or bounds a parameter of interest. One generally expresses this faith statistically by affixing to each interval a confidence coefficient, or confidence probability, which can be written as $1-\epsilon$, where $\epsilon = p/100$ for $0 \leq p \leq 100$, and p is usually taken to be a very small number (**1**, **3**, **6**, **8**). For example, if $p = 5$ the confidence coefficient would be .95, and one would refer to the interval with which this coefficient is associated as the 95% confidence interval.

The confidence coefficient is frequently interpreted in the following manner: If one were to draw samples of size K from a population of N elements (K naturally being $< N$) and from each sample establish a 95% confidence interval on some specified parameter of the population, then in the long run about 95% of the totality of these intervals would actually contain the parameter of interest, and approximately 100 ϵ%, or 5%, of them would not (**3**). This interpretation is correct, but of course assumes $\binom{N}{K}$ to be a rather large number.[1]

SIGNIFICANCE AND SIGNIFICANCE LEVELS

Significance, as contrasted to confidence, is given to the testing of hypotheses. Here one makes a statement, i.e., states an hypothesis, which will hereafter in this discussion be represented as H, that may be either true or false and then takes action on this H by accepting or rejecting it. Clearly, any one of the following actions is a likely outcome as a result of testing an H: (*a*) rejection of a false H; (*b*) acceptance of a true H; (*c*) rejection of a true H; or (*d*) acceptance of a false H. It is quite evident that actions (*a*) and (*b*) are desirable, while (*c*) and (*d*) carry the connotation of committing an error—*c* being the familiar Type I error or an error of the first kind, while *d* is called a Type II error or an error of the second kind (**4**, **8**).

When one tests an H, the probability that he will take action *c* is defined as the significance level, which we will represent as α (**8**). Although α is generally of the same order of

[1] The notation $\binom{N}{K}$ is used here as a combinational symbol to indicate the number of ways that K objects can be selected from N.

magnitude as ϵ, α and ϵ differ in the amount of information which they convey, for while ϵ completely tells all there is to know about "being wrong" in interval estimation, α only gives information about a very particular type of error, i.e. the action described by c. To emphasize the contrast made here between α and ϵ, one merely needs to examine the other type of error that can be made in the test of an H.

For this purpose, let β represent the probability that action d is taken, i.e. a Type II error is committed; then, by definition, $1-\beta$ is known as the power of the statistical test or the probability that action a will occur (**4, 8**). Although texts in psychological statistics do not seem to place a great deal of emphasis upon the power of a test, power is the basic concept responsible for one's employing statistical tests as a basis for taking action on an H. If this were not so, to test an H at the 5% level of significance, one could simply draw from a box of 100 beads—95 white and 5 red—a bead at random and adopt the convention that he would reject the H whenever a red bead appeared. With such a test, one can readily see that not only α but also $1-\beta$ always equals .05, or $\beta = .95$. It is this large value of β that precludes one's employing the bead-box test. For an excellent discussion of β and its relation to the alternative H against which one might be testing, the reader is referred to Dixon and Massey (**4**, pp. 244–261).

Summary and Discussion

The admixing of the concepts of confidence and significance has become so prevalent in the psychological literature that one typically reads statements, in the reports of psychological research, indicating that certain experimental results were significant at, say, the 5% "level of confidence."

It may be that this confusion arises from the fact that one can utilize a confidence interval as a significance test (e.g., see **5**, p. 241), and in doing so may hastily, but incorrectly, conclude that there is no difference between the two concepts.

Inasmuch as explicit terminology is needed to convey the probabilities of committing statistical errors in the respective areas of interval estimation and testing of hypotheses, the concept of confidence should never be associated with the statistical test of an H regardless of the nature of the test being employed.

REFERENCES

1. Anderson, R. L., & Bancroft, T. A. *Statistical theory in research.* New York: McGraw-Hill, 1952.
2. Chandler, R. E. A review of Guilford's *Fundamental statistics in psychology and education. (3rd ed.) Personnel Psychol.*, 1957, **10**, 272–273.
3. Cramér, H. *Mathematical methods of statistics.* Princeton: University Press, 1951.
4. Dixon, W. J., & Massey, F. J., Jr. *Introduction to statistical analysis.* (2nd ed.) New York: McGraw-Hill, 1957.
5. Edwards, A. L. *Statistical methods for the behavioral sciences.* New York: Rinehart, 1954.
6. Hoel, P. G. *Introduction to mathematical statistics.* (2nd ed.) New York: Wiley, 1954.
7. Milton, T. E. A review of Edwards' *Statistical methods for the behavioral sciences. J. Amer. statist. Ass.*, 1956, **51**, 382.
8. Mood, A. M. *Introduction to the theory of statistics.* New York: McGraw-Hill, 1950.
9. Walker, H. M. A review of Adams' *Basic statistical concepts. Educ. psychol. Measmt*, 1956, **16**, 554–557.

Received May 20, 1957.

ARE THERE GAPS IN YOUR FILES OF APA JOURNALS?

SALE of BACK ISSUES

Here's how this . . .

The American Psychological Association announces a sale during the period October 1957 through March 1958. Of the following journals, all available issues in the volume for the years preceding 1961 will be offered at a price of only 50¢ (foreign, 60¢) per issue:

American Psychologist
Journal of Abnormal & Social Psychology
Journal of Applied Psychology
Journal of Comparative & Physiological Psychology (1947-1960 only)
Journal of Consulting Psychology
Journal of Experimental Psychology
Psychological Abstracts
Psychological Bulletin
Psychological Index (a few complete volumes, some shopworn)
Psychological Monographs
Psychological Review

Not all issues in all volumes are available. But—ORDER NOW before more back issues go out of print. From our available stock we will complete as much of your order as possible at this reduced price and for this limited period.

Delivery: 6 to 8 weeks — No dealer or quantity discounts

After this sale, for the years preceding 1948, journals will be available only on microfilm and microcard.

Order from:
American Psychological Association
Department BB
1333 Sixteenth Street, N.W.
Washington 6, D. C.

GEORGE BANTA COMPANY, INC., MENASHA, WISCONSIN, U.S.A.